

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:11:45 ; Search time 22 Seconds
(without alignments)
4859.092 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10490
Sequence: 1 MLKFKYGARNPLDAGAEPI.....QLNGEIROQVEKSVLRDTC 2053

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7491.5	71.4	1597	1 CTRO_MOUSE	P49025 mus musculus
2	6223.5	59.3	1286	1 CTRO_HUMAN	Q14578 homo sapien
3	862	531	1	1 DMK_MOUSE	P54265 mus musculus
4	839.5	8.0	639	1 DMK_HUMAN	Q09013 homo sapien
5	698	6.7	756	1 CBK1_YEAST	P53894 saccharomyc
6	691	6.6	469	1 ORB6_SCHPO	Q13310 schizosacch
7	630	6.0	598	1 COT1_NEUCR	P38579 neurospora
8	590.5	5.6	1411	1 EEAL_HUMAN	Q15075 homo sapien
9	575.5	5.5	2238	1 GOA4_MOUSE	Q91vw5 mus musculu
10	561	5.3	2442	1 CEP2_HUMAN	Q9bv73 homo sapien
11	551	5.3	2230	1 GOA4_HUMAN	Q13439 homo sapien
12	540.5	5.2	2116	1 MYS2_DICDI	P08799 dictyosteli
13	540	5.1	444	1 KRAC_DICDI	P54644 dictyosteli
14	540	5.1	607	1 KATB_SCHPO	Q09398 schizosacch
15	538	5.1	1976	1 MYHA_BOVIN	Q27951 bos taurus
16	536	5.1	1976	1 REST_CHICK	Q42184 gallus gall
17	535.5	5.1	1433	1 USO1_YEAST	P25386 saccharomyc
18	534	5.1	1790	1 MYB1_HUMAN	P35580 homo sapien
19	534	5.1	1976	1 GOB1_HUMAN	Q14789 homo sapien
20	533.5	5.1	3259	1 TRH1_RABIT	P37709 oryctolagus
21	532	5.1	1407	1 MLPI_YEAST	Q02455 saccharomyc
22	531.5	5.1	1875	1 TPR_HUMAN	P12270 homo sapien
23	529.5	5.0	2349	1 MYSB_CAEEL	P02566 caenorhabdi
24	529	5.0	1966	1 CENE_HUMAN	Q02224 homo sapien
25	528.5	5.0	2663	1 C190_DROME	Q9vj05 drosophila
26	526.5	5.0	1690	1 MYSN_DROME	Q93323 drosophila
27	526.5	5.0	2017	1 DBF2_YEAST	P32328 saccharomyc
28	524.5	5.0	564	1 PLE1_HUMAN	Q15149 homo sapien
29	524.5	5.0	4684	1 PLE1_HUMAN	P35579 homo sapien
30	524	5.0	1960	1 MYH9_HUMAN	P12844 caenorhabdi
31	523.5	5.0	1969	1 MYSB_CAEEL	P35749 homo sapien
32	521	5.0	1972	1 MYHB_HUMAN	P35748 oryctolagus
33	520.5	5.0	1972	1 MYHB_RABIT	

ALIGNMENTS

RESULT 1

CTRO_MOUSE STANDARD; PRT; 1597 AA.
AC P49025;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Citron protein (Rho-interacting, serine/threonine kinase 21).
GN CIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128238; PubMed=8543060;
RA Madanle P., Furuyaashiki T., Reid T., Ishizaki T., Watanabe G.,
RA Morii N., Narumiya S.;
RT "A novel partner for the GTP-bound forms of rho and rac.";
RL FEBS-Lett-377:243-248(1995).
CC FUNCTION: Putative RHO/RAC effector that binds to the GTP-bound
CC forms of RHO and RAC1. It probably binds p21 with a tighter
CC specificity in vivo.
CC SUBUNIT: Homodimer (Probable).
CC TISSUE SPECIFICITY: A major signal was observed in testis and
CC brain, but it was also detected in thymus, spleen, kidney, heart
CC and lung.
CC SIMILARITY: Contains 1 CNH domain.
CC SIMILARITY: Contains 1 PH domain.
CC SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
CC binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; U39904; AAC52341.1; -.
CC PIR; S68420; S68420.
CC MGD; MGI:105313; Cit.
CC InterPro; IPR001180; Citron.
CC InterPro; IPR002219; DAG_PE-bind.
CC InterPro; IPR001849; PH.
CC InterPro; IPR000861; REM_repeat.
CC Pfam; PF00780; CNH; 1.
CC Pfam; PF00130; DAG_PE-bind; 1.
CC Pfam; PF00169; PH; 1.
CC SMART; SM00109; Cl; 1.
CC SMART; SM00036; CNH; 1.
CC SMART; SM00233; PH; 1.
CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
CC PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.

Q9j155 cricetus
Q9Y623 homo sapien
P11055 homo sapien
P24733 aequipecten
P10567 gallus gall
Q02566 mus musculus
P22204 saccharomyc
Q90339 cyprinus ca
P30427 rattus norv
P13533 homo sapien
P53739 saccharomyc
P13539 mesocricetu

DR PROSITE; PS0003; PH DOMAIN; 1.
KW Coiled coil; Phorbol-ester binding; SH3-binding.
FT DOMAIN 1 845 COILED COIL (POTENTIAL).
FT DOMAIN 674 870 RHO/RAC BINDING.
FT DOMAIN 818 821 POLY-LYS.
FT DOMAIN 931 979 PHORBOL-ESTER AND DAG BINDING.
FT DOMAIN 1011 1131 PH.
FT DOMAIN 1160 1457 CNH.
FT SITE 1521 1526 SH3-BINDING (POTENTIAL).
SQ SEQUENCE 1597 AA; 183448 MW; 787286C2305676DA CRC64;

Query Match 71.4%; Score 7491.5; DB 1; Length 1597;
Best Local Similarity 93.1%; Pred. No. 8.8e-224;
Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;

QY 468 MEQEMTLHRRVSEVAVLSQKEVELKASQVVEEMRLMNMQLLEEDLVSAARRSDLYES 527
DB 9 MEQEMTLHRRVSEVAVLSQKEVELKASQVVEEMRLMNMQLLEEDLVSAARRSDLYES 68
QY 528 VSQEDDKALQLLHDIREQSRKLQEIKEQEVQAOVEEMRLMNMQLLEEDLVSAARRSDLYES 587
DB 69 VSQEDDKALQLLHDIREQSRKLQEIKEQEVQAOVEEMRLMNMQLLEEDLVSAARRSDLYES 128
QY 588 ELRESRLAAEFKFKKATECHQKLLKAKDQKQPEVGEYAKLEKINAEQQLKIQBLQKLEK 647
DB 129 ELRESRLAAEFKFKKATECHQKLLKAKDQKQPEVGEYAKLEKINAEQQLKIQBLQKLEK 188
QY 648 A-----AKRAERLEKLNRRDSSEGIRKLVAEERRHSLNKKVL 691
DB 189 AVKASTFATELLQNIQAKAPARRELEKLNRRDSSEGIRKLVAEERRHSLNKKVL 248
QY 692 ETWERENRLKDDIQTKSQIQQWADKILELEKHEKREAVSAQHLVHLKQEQHYEEKI 751
DB 249 ETWERENRLKDDIQTKSQIQQWADKILELEKHEKREAVSAQHLVHLKQEQHYEEKI 308
QY 752 KVLNDQIKDLADKETLNNMQHSEAEKILSEKAMINAMDSKIRSEQRIVELS 811
DB 309 KVLNDQIKDLADKETLNNMQHSEAEKILSEKAMINAMDSKIRSEQRIVELS 368
QY 812 EANKLAANSLSFTQNNKAQEMISELRQKQFYLETQAGKLEAQNFKLEQEKISHQDH 871
DB 369 EANKLAANSLSFTQNNKAQEMISELRQKQFYLETQAGKLEAQNFKLEQEKISHQDH 428
QY 872 SDKNRLLELTRUREVSLHEEQKLEKQLTLEQLSLQERESQTLQAARAALESQLR 931
DB 429 SDKNRLLELTRUREVSLHEEQKLEKQLTLEQLSLQERESQTLQAARAALESQLR 488
QY 932 QAKTELEETTAEEETQALTAHRDIQKFDALRNSCTVITDLEBQLNQLTDEDAELNN 991
DB 489 QAKTELEETTAEEETQALTAHRDIQKFDALRNSCTVITDLEBQLNQLTDEDAELNN 548
QY 992 QNPLYSKQLDEAGANDEIVQLASEVDHLRREITEREMQLTSQKQTMELKTTCTMLBEQ 1051
DB 549 QNPLYSKQLDEAGANDEIVQLASEVDHLRREITEREMQLTSQKQTMELKTTCTMLBEQ 608
QY 1052 VMLEALNDELLEKQEWAEARVSLGDEKQSECEVRELRQMLDTEKQSRARADQRTES 1111
DB 609 VLDLEALNDELLEKQEWAEARVSLGDEKQSECEVRELRQMLDTEKQSRARADQRTES 668
QY 1112 QVVELAVVHEKHAIIALQALKEQKKAESISDKLNDLEKHAMLENNARSLOQKLETE 1171
DB 669 QVVELAVVHEKHAIIALQALKEQKKAESISDKLNDLEKHAMLENNARSLOQKLETE 728
QY 1172 RELKQRLLEEQAKLQOQMDLQKXHIIFRLTQGLQEAALDRADLKLKTRSDLEYOLENIQVLY 1231
DB 729 RELKQRLLEEQAKLQOQMDLQKXHIIFRLTQGLQEAALDRADLKLKTRSDLEYOLENIQVLY 788
QY 1232 SHEKVRMEGTISQTKLIDFLQAKMDQPAKKKGLFSRRKEDPALTPQVLPQYNELKAL 1291
DB 789 SHEKVRMEGTISQTKLIDFLQAKMDQPAKKK-----VPLQYNELKAL 833
QY 1292 EKEKARCAELEALQKTRIELSAREEAHAKATDHPHPSTPATARQOIAMSAIVRSPEH 1351

DB 834 EKEKARCAELEALQKTRIELSAREEAHAKATDHPHPSTPATARQOIAMSAIVRSPEH 893
QY 1352 QPSAMSLIAPPSSRRKESSTPEEFSRRLKERMHNIPIHRFNVGLNMRATKCAVCLDTVHF 1411
DB 894 QPSAMSLIAPPSSRRKESSTPEEFSRRLKERMHNIPIHRFNVGLNMRATKCAVCLDTVHF 953
QY 1412 GRQASKCLECOVMCHPKKSTCLPATCGLPABYATHFEAFCDKQNSPGLQTKPESSSLH 1471
DB 954 GRQASKCLECOVMCHPKKSTCLPATCGLPABYATHFEAFCDKQNSPGLQTKPESSSLH 1013
QY 1472 LEGMKVPRNNKRGQGWDRKIIVLEGSKVLIYDNEAREAGQREVEFEFELCLPDGDSIH 1531
DB 1014 LEGMKVPRNNKRGQGWDRKIIVLEGSKVLIYDNEAREAGQREVEFEFELCLPDGDSIH 1073
QY 1532 GAVGASELANATAKA----- 1545
DB 1074 GAVGASELANATAKADVPYILKQESHPTTCWPGRTLYLLAPSPDKORWVTALESVAVG 1133
QY 1546 ---EKAPADAKLLGNSLLKLEGDRDLMNCTLPFSQVVLVGTTEGLYALNVLKNSLTH 1601
DB 1134 RVSREKAEADAKLLGNSLLKLEGDRDLMNCTLPFSQVVLVGTTEGLYALNVLKNSLTH 1193
QY 1602 VPGIGAVPQIYIIKDLKLLMAGEERALCLVDVKKYKQSLAQSHLPAQDPDISNIPEAV 1661
DB 1194 IPGIGAVPQIYIIKDLKLLMAGEERALCLVDVKKYKQSLAQSHLPAQDPDISNIPEAV 1253
QY 1662 KGCHLFGAGKLENGLCICAAMPKSVILRYNENLSKYCIKKEITSTSPCIIHTNYSIL 1721
DB 1254 KGCHLFGAGKLENGLCICAAMPKSVILRYNENLSKYCIKKEITSTSPCIIHTNYSIL 1313
QY 1722 IGTNKFYEIDMKQYTLLEFELDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBYLLCFH 1781
DB 1314 IGTNKFYEIDMKQYTLLEFELDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBYLLCFH 1373
QY 1782 EFGVFDVSYGPRSRDDELKWSRLPLAFAYREPILFVTHNSLEVIEIQARSSAGTTPARAY 1841
DB 1374 EFGVFDVSYGPRSRDDELKWSRLPLAFAYREPILFVTHNSLEVIEIQARSSAGTTPARAY 1433
QY 1842 LDIPNRYLGAISSGAIYLAASSYQDKLRVICCKGNLVKSGTEHHRGPGSTSRSSPNKRG 1901
DB 1434 LEIPNRYLGAISSGAIYLAASSYQDKLRVICCKGNLVKSGTEHHRGPGSTSRSSPNKRG 1493
QY 1902 PPTNEHITKRVASSPAPPGSPHPRPSTPHRY--REGHTELRKXSGRPLEREKSPG 1959
DB 1494 PPTNEHITKRVASSPAPPGSPHPRPSTPHRYRDREGHTELRKXSGRPLEREKSPG 1553
QY 1960 RMLSTRERSPGRLFFEDSSRGRIPAGAVRTPLSQVNVKGRQSA 2002
DB 1554 RMLSTRERSPGRLFFEDSSRGRIPAGAVRTPLSQVNVKGRQSA 1596

RESULT 2
CTRO HUMAN
ID CTRO HUMAN STANDARD; PRT; 1286 AA.
AC 014578; 09UP27;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Citron protein (Rho-interacting, serine/threonine kinase 21)
DE (Fragment).
DE CIT OR STK21 OR KIAA0949.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Connell M., Goela D., Harper M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE OF 347-1286 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99246063; PubMed=10231032;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroseawa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:63-70(1999).
CC -!- FUNCTION: Putative RHO/RAC effector that binds to the GTP-bound
CC forms of RHO and RAC1. It probably binds p21 with a tighter
CC specificity in vivo (By similarity).
CC -!- SIMILARITY: Contains 1 CNH domain.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
CC binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AC002563; AAB71327.1; -;
CC EMBL; AB023166; BAA76793.1; -;
CC Genew; HGNC:1985; CIT.
CC MIM; 605629; -;
CC InterPro; IPR001180; Citron.
CC InterPro; IPR002219; DAG_PE-bind.
CC InterPro; IPR001849; PH.
CC InterPro; IPR000861; REM_repeat.
CC Pfam; PF00780; CNH; 1.
CC Pfam; PF00130; DAG_PE-bind; 1.
CC Pfam; PF00169; PH; 1.
CC SMART; SM00109; C1; 1.
CC SMART; SM00036; CNH; 1.
CC SMART; SM00233; PH; 1.
CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
CC PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
CC PROSITE; PS00003; PH_DOMAIN_1;
KW Coiled coil; Phorbol-ester binding; SH3-binding.
FT NON TER 1
FT DOMAIN <1 558 COILED COIL (POTENTIAL).
FT DOMAIN 365 561 RHO/RAC BINDING.
FT DOMAIN 509 512 POLY-LYS.
FT DOMAIN 622 670 PHORBOL-ESTER AND DAG BINDING.
FT DOMAIN 702 822 PH.
FT DOMAIN 851 1148 CNH.
FT SITE 1212 1217 SH3-BINDING (POTENTIAL).
SQ SEQUENCE 1286 AA; 146506 MW; 498101F79EA75E85 CRC64;
Query Match 59.3%; Score 6223.5; DB 1; Length 1286;
Best Local Similarity 94.7%; Pred. No. 6.8e-185;
Matches 1231; Conservative 1; Mismatches 3; Indels 65; Gaps 2;
QY 753 VLDNQIKKLADKETLENMQRHEEBAHKGKILSEKAMINAMDSKIRSLQRIVELSE 812
DB 1 VLDNQIKKLADKETLENMQRHEEBAHKGKILSEKAMINAMDSKIRSLQRIVELSE 60
QY 813 ANKLAANSSIFLTKNNKAQEMISELRQOKFYLETQAGKLEAQNRLKEOLEKISHQDS 872
DB 61 ANKLAANSSIFLTKNNKAQEMISELRQOKFYLETQAGKLEAQNRLKEOLEKISHQDS 120
QY 873 DKNRLELETRLRREVSLHEEQKLEKRLQTELQSLQRESQLTALQARAALLESQLRQ 932
DB 121 DKNRLELETRLRREVSLHEEQKLEKRLQTELQSLQRESQLTALQARAALLESQLRQ 180
QY 933 AKTELETTAAEEETQALTAHDETCRPFDAIRNSCTVITDLEEQNLQITDNEALNNQ 992
DB 181 AKTELETTAAEEETQALTAHDETCRPFDAIRNSCTVITDLEEQNLQITDNEALNNQ 240
QY 993 NFVLSKQDLEASGANDIEIVLRSEVDHLRREITERENQOLTSQKTWEALKTCTMLEEQV 1052
DB 241 NFVLSKQDLEASGANDIEIVLRSEVDHLRREITERENQOLTSQKTWEALKTCTMLEEQV 300

RESULT 3

DMK_MOUSE

ID DMK_MOUSE STANDARD; PRT; 631 AA.


```
SQ SEQUENCE 631 AA; 69601 MW; 5EB800A37EA81DF4 CRC64;
Query Match 8.2%; Score 862; DB 1; Length 631;
Best Local Similarity 36.6%; Pred. No. 2.3e-20;
Matches 199; Conservative 104; Mismatches 191; Indels 50; Gaps 13;

QY 46 LSGREGIDALFVLFECSPALMKIKHVNFKYSDTIAELQELQPSAKDFEVSRLVGC 105
DB 20 LGLEPFLDLLGVHQLGASHLAQDKYVADFQWVEPIAARLKEVRLQDRDDFEILKVIGR 79
QY 106 GHFAEVQVVRKATGDIYAKVMKKKALLAQSOVFFEEERNILSRSTSPWLPQLOAFQ 165
DB 80 GAFSEVAVVMKQTGQVYAKIMKMKMDMLKRGVSCFEEERDVLVKGDRMTQLHFAFQ 139
QY 166 DKNHLYMEXYQPGDLSILARYEDQDENLIQFYLAELIIVAVSHVLMGVVHRDIKPE 225
DB 140 DENILYLVMEYVYGGDLTLKSKFGERIPAEWARFVLAIVVAIDSVHRLGVVHRDIKPD 199
QY 226 NILVDRTHIKLVDFGSAAMNSNMKNVNAKLPIGTPDYMAPEVLTVMNGDGKGYGLDC 284
DB 200 NILDCRGHRLADFGSLKLPQDGMVRSILVAVGTPDYLSPILQAVGGPGAGSYGPEC 259
QY 285 DMSVGVIAIYEMTYGRSPAEGTSARTFNINMNFQFLKFPDDPKVSSDFDLILQSLLC 343
DB 260 DMWALGVFAYEMPYGGTTPYADSTAYAKIYHREHLSLPLADTVVPEEAQDLIRGLLC 319
QY 344 GQKERLKFEGL--LCCHPFESKIDWNINRNSPPFVPTLKSDDDTSNFD--EPEKSNMVS 398
DB 320 PAEIRLGRGAGDFQKHFFFGDLWGLSDSVPPFTDFEGATDTCNFDVVEDRLTAMVS 379
QY 399 SPCQLSPSGFS--GEELPFVGFYSKALGILGRSESVVGLDSPAKTSSMEKILLIKS 455
DB 380 GGGETLSMDQEDMPLGVRLEFVGYSY--CCNAFRDNQV-----PDPTFMLEALQLPV 430
QY 456 KEIQ--DSQDKCHKMEQ-----EMTLHRRVSEVAVLQKVELKASQTSRLLEQDLA 508
DB 431 SDLQGLDLQPPVPPQVQAEADLVAVPAPVAEATTVTLQQLQ-----EALSEEV 482
QY 509 TYITECSLKSRLSEARMEVQSDKALQLLHDIREQSKRLQEIKEQYQAOVVEMLMM 568
DB 483 TR----QSLRELEAIR-TANQNFSSQLQ-----EAEVRNRDLEAHVRLQIERM 526
QY 569 NQLE 572
DB 527 EMLQ 530

RESULT 4
DMK_HUMAN
ID DMK_HUMAN STANDARD; PRT; 639 AA.
AC Q09013; Q16205;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myotonin-protein kinase [EC 2.7.1.-] (Myotonic dystrophy protein
DE kinase) (MDPK) (DM-kinase) (DMK) (DMPK) (MT-PK).
GN DMPK OR MDPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92188202; PubMed=1546326;
RA Fu Y.-H., Pizzuti A., Fenwick R.G. Jr., King J., Rajnarayan S.,
RA Dunne P.W., Dubel J., Nasser G.A., Ashizawa T., de Jong P.J.,
RA Wieringa B., Korneluk R., Perryman M.B., Epstein H.F., Caskey C.T.;
RT "An unstable triplet repeat in a gene related to myotonic muscular
RT dystrophy."
RL Science 255:1256-1258 (1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 5; 6; 7 AND 8).
RX MEDLINE=93227030; PubMed=8469976;
Fu Y.-H., Friedman D.L., Richards S., Pearlman J.A., Gibbs R.A.,
Pizzuti A., Ashizawa T., Perryman M.B., Scariato G., Fenwick R.G. Jr.,
Caskey C.T.;
"Decreased expression of myotonin-protein kinase messenger RNA and
protein in adult form of myotonic dystrophy."
Science 260:235-238 (1993).
[3]
RP SEQUENCE FROM N.A. (ISOFORMS 9; 10; 11 AND 12).
RX MEDLINE=93271990; PubMed=8499920;
RA Mahadevan M.S., Amemiya C., Jansen G., Sabourin L., Baird S.,
RA Neville C.E., Wormskamp N., Segers B., Batzer M., Lamerdin J.,
RA de Jong P.J., Wieringa B., Korneluk R.G.,
RT "Structure and genomic sequence of the myotonic dystrophy (DM kinase)
RT gene."
Hum. Mol. Genet. 2:299-304 (1993).
[4]
RP SEQUENCE FROM N.A. (ISOFORM 11).
RC TISSUE=Brain, and Muscle;
RX MEDLINE=94140369; PubMed=7905855;
RA Shaw D.J., McCurrach M., Rundle S.A., Harley H.G., Crow S.R., Sohn R.,
RA Thirion J.-P., Hamshire M.G., Buckler A.J., Harper P.S., Housman D.E.,
RA Brook J.D.;
"Genomic organization and transcriptional units at the myotonic
RT dystrophy locus."
Genomics 18:673-679 (1993).
[5]
RP SEQUENCE FROM N.A. (ISOFORM 11).
RC TISSUE=Muscle;
RX MEDLINE=94357271; PubMed=8076586;
RA Sasagawa N., Sorimachi H., Maruyama K., Arakata K., Ishiura S.,
RA Suzuki K.;
"Expression of a novel human myotonin protein kinase (MCPK) cDNA clone
RT which encodes a protein with a thymopoleitin-like domain in COS
RT cells."
FEBS Lett. 351:22-26 (1994).
[6]
RP SEQUENCE OF 24-639 FROM N.A. (ISOFORMS 9 AND 11).
RC TISSUE=Brain, and Heart;
RX MEDLINE=93251003; PubMed=1302022;
RA Jansen G., Mahadevan M., Amemiya C., Wormskamp N., Segers B.,
RA Hendriks W., O'Hoy K., Baird S., Sabourin L., Lennon G., Jap P.L.,
RA Iles D., Coerwinkel M., Hofker M., Carrano A.V., de Jong P.J.,
RA Korneluk R.G., Wieringa B.;
"Characterization of the myotonic dystrophy region predicts multiple
RT protein isoform-encoding mRNAs."
Nat. Genet. 1:261-266 (1992).
[7]
RP SEQUENCE OF 52-639 FROM N.A. (ISOFORM 13).
RX MEDLINE=92154692; PubMed=1310900;
RA Brook J.D., McCurrach M., Harley H.G., Buckler A.J., Church D.,
RA Aburatani H., Hunter K., Stanton V.P., Thirion J.-P., Hudson T.,
RA Sohn R., Zemelman B., Snell R.G., Rundle S.A., Crow S., Davies J.,
RA Shelbourne P., Buxton J., Jones C., Juvonen V., Johnson K.,
RA Harper P.S., Shaw D.J., Housman D.E.;
"Characterization of myotonic dystrophy: expansion of a trinucleotide
RT (CTG) repeat at the 3' end of a transcript encoding a protein kinase
RT family member."
Cell 68:799-808 (1992).
[8]
RP SEQUENCE OF 539-575 FROM N.A. (ISOFORM 14).
RC TISSUE=Brain;
RX MEDLINE=96063635; PubMed=7488138;
RA Gennarelli M., Lucarelli M., Zelano G., Pizzuti A., Novelli G.,
RA Dallapiccola B.;
"Differeant expression of the myotonin protein kinase gene in discrete
RT areas of human brain."
Biochem. Biophys. Res. Commun. 216:489-494 (1995).
RL Biochem. Biophys. Res. Commun. 216:489-494 (1995).
CC -!- FUNCTION: May play a role in intracellular communication.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=12;
CC Name=1;
CC IsoId=Q09013-1; Sequence=Displayed;
CC Name=2;
```


Db 528 RLOQERNELLOAE 540

RESULT 5

CBK1_YEAST STANDARD; PRT; 756 AA.

AC P53894; 1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine/threonine-protein kinase CBK1 (EC 2.7.1.1.) (cell wall

DE biosynthesis kinase)

GN CBK1 OR YN1161W OR Nt727.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / FY1679;

RX MEDLINE=96287653; PubMed=8686380;

RA Nasr F., Becam A.-M., Herbert C.J.;

RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals

RT 24 complete open reading frames; 18 correspond to new genes, one of

RT which encodes a protein similar to the human myotonic dystrophy

RT kinase.";

RL Yeast 12:169-175 (1996).

RN [2]

RP INTERACTION WITH PAG1/TAO3.

RX MEDLINE=21843681; PubMed=11854408;

RA Du L.L., Novick P.;

RT "Pag1p, a novel protein associated with protein kinase Cbx1p, is

RT required for cell morphogenesis and proliferation in Saccharomycetes

RT cerevisiae";

RL Mol. Biol. Cell 13:503-514 (2002).

CC -1- FUNCTION: Protein kinase that seems to play a role in thr

CC regulation of cell morphogenesis and proliferation.

CC -1- SUBUNIT: Associates with PAG1/TAO3.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC COT1 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X92517; CAA63278.1; -.

DR EMBL; Z71437; CAA96048.1; -.

DR PIR; S60966; S60966.

DR GERMOnline; 143167; -.

DR SGD; S0005105; CBK1.

DR GO; GO:0005935; C:bud neck; IDA.

DR GO; GO:0005737; C:cytoplasm; IDA.

DR GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0000753; P:cellular morphogenesis during conjugation w. . . ; IDA.

DR GO; GO:0030012; P:establishment and/or maintenance of cell po. . . ; IMP.

DR GO; GO:0007096; P:regulation of exit from mitosis; IMP.

DR InterPro; IPR000961; Pkinase_C

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR008271; Ser_Thr_Pkin_AS.

DR InterPro; IPR002290; Ser_Thr_Pkinase.

DR Pfam; PF00069; pkinase_1.

DR Pfam; PF00433; pkinase_C; 1.

DR ProDom; PD000001; Prot_kinase; 2.

DR SMART; SM00133; S_TK_X; 1.

DR SMART; SM00220; S_TK; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS00108; PROTEIN KINASE ST; 1.

DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.

DR Transferase; Serine/threonine-protein kinase; ATP-binding.

PT DOMAIN 23 34 POLY-GLN.

PT DOMAIN 160 164 POLY-SER.

PT DOMAIN 214 232 POLY-GLN.

PT DOMAIN 235 250 POLY-GLN.

PT DOMAIN 352 672 PROTEIN KINASE.

PT NP BIND 358 366 ATP (BY SIMILARITY).

PT BINDING 381 381 ATP (BY SIMILARITY).

PT ACT SITE 475 475 BY SIMILARITY.

SQ SEQUENCE 756 AA; 86946 MW; 87EBDCD2C3C96EE11 CRC64;

Query Match 6.7%; Score 698; DB 1; Length 756;

Best Local Similarity 32.5%; Pred. No. 3e-15;

Matches 162; Conservative 83; Mismatches 159; Indels 94; Gaps 10;

QY 4 PKYCARNP--LDAGAAPPIASRASRLNLFQGGKPPFTQQQMSPLSREGILDALFVLFEE 61

DB 265 YMYFERPDLTKGTQDAKAVLKIKENFYQSSVKYAIERNRVERVELETETSHWSEER 324

QY 62 CSQPALMKIKHVSNFVRKYSDTIAELOPQSAKDFEVRSLVGGCHFAEQVVRKATGD 121

DB 325 KSRQLSSLGKESQFLR-----LRRTRLSLEDFHTVKVIGKGAFAFEVRLVQKDTGK 376

QY 122 IYANKMKKALLAQEQVSPFEERNILSRSTSPWIPQLOVAFODKXHLVLMEEYQPGGD 181

DB 377 IYANKTLKSEMYKKDQLAHVKAERDVLAGSDSPWVSLYYSFQDAQYLYLIMEFLPGGD 436

QY 182 LLSLINRYEDQDENLIQFVLAELIYVHSLVHMGYVHRDIKPNILVDRTHGIKLVDFG 241

DB 437 LMTMLIRWQ-LFTEDVTRFYMAECILAEITHIKLGFHRIKPDNILDIRGHIKLSDFG 495

QY 242 -----SAAKMNSN----- 249

DB 496 LSTGFHKTHDSNYKKLLQDQEAENG:SKPGTYNANTTDTANKRQTWVVDISISITMSNRQ 555

QY 250 -----KMNNAKLPFGTDPYMAPEVLTVMNGDGKGTGYGLDCDWWSVGVYIAYEMYGSR 301

DB 556 QIQWKRKSRRLWAYSTVGTPTYIAPEIFYQG-----YQECDDWSLGAIMYECLIGWP 609

QY 302 PPAEGTSARTSNMNPORFLKFPDDPKVSSDFLDLIQSLCGQKERLKFEQ-----LCCH 357

DB 610 PFCSETPQETRYKIMNFETLQFPDDIHISYEAEELIRLLTHADQRLGRHGGADEIKSH 669

QY 358 PFESKIDMNNIRNSPPRPVPTLKSDDTSNDFEPEKSNWSVSSPC-----QLSPSG 408

DB 670 PFFRGVDWNTIRQVEAPYIPKLSITDRFFPTDELEN-VFDSPAMAQAQKREQMTQGG 728

QY 409 FSG-----EELPFVGFYSYK 423

DB 729 GSAPVKEDLPFGITYSR 746

RESULT 6

ORB6_SCHPO

ID ORB6 SCHPO STANDARD; PRT; 469 AA.

AC O13310;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine/threonine-protein kinase orf6 (EC 2.7.1.37).

GN ORB6 OR SPAC821.12.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98301608; PubMed=9636183;

RA Verde F., Wiley D.J., Nurse P.;

RT "Fission yeast orf6, a ser/thr protein kinase related to mammalian

RT rho kinase and myotonic dystrophy kinase, is required for maintenance

RT of cell polarity and coordinates cell morphogenesis with the cell

RT cycle.";

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X97657; CAA66253.1; -
DR PIR; S22711; S22711; -
DR PIR; T47254; T47254; -
DR HSSP; P05132; 1BKX.
DR InterPro; IPR000961; Pkinase C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser Thr pkinase.
DR Pfam; PFC0069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative initiation.
FT CHAIN 1 598
FT SERINE/THREONINE-PROTEIN KINASE COT-1,
FT ISOFORM LONG.
FT CHAIN 119 598
FT SERINE/THREONINE-PROTEIN KINASE COT-1,
FT ISOFORM SHORT.
FT INIT MET 119 598
FT FOR ISOFORM SHORT.
FT DOMAIN 214 518
FT PROTEIN KINASE.
FT NP BIND 220 228
FT ATP (BY SIMILARITY).
FT BINDING 243 243
FT ATP (BY SIMILARITY).
FT ACT SITE 337 337
FT BY SIMILARITY.
SQ SEQUENCE 598 AA; 68017 MW; 13B9FBAF3859893 CRC64;

Query Match 6.0%; Score 630; DB 1; Length 598;
Best Local Similarity 34.8%; Pred. No. 2.9e-13;
Matches 134; Conservative 81; Mismatches 107; Indels 66; Gaps 9;

QY 95 KDFEVSILGCGHFAEVQVVRKATGDIVAMKMKKALLAQVQFFFEERNILSRSTS 154
Db 212 ENVTQTIKIGKGAFAVEKLVQKADGVAMKSLIKTEFMKQDLAHVRAERDILAESDS 271
QY 155 PWTPQLOAFQDKNHLYLMEYQPGDLSLLNRYEDQDENLIQYLAELILAVSHVL 214
Db 272 PWYKLYTTQDANFLYMLMEFLFGDGLMTLIKYE-IFSEDITREYIAEIVLAIDAVHK 330
QY 215 MGVVHRDIKPNILVDTGHKLVDFG----- 241
Db 331 LGFIHRDIKPNILVDTGHKLVDFG----- 390
QY 242 -----SAAKNS---NKVMNAKLPITGPDYMAPEVLTVNMGKGTGGLDCDWNSV 289
Db 391 IDQINLTVSNRAQINDWRSRRLMAYSTVGTPIAIEFT-----GHG-YSPDCDWNSL 444
QY 290 GVTAYEMVYGRSFAEGTSARTENNTNFORFLKFPDDPKVSSDFDLDTQSLCGOKEL 349
Db 445 GTMFECLVGPWPFCAEDSDHTYKLVNMHSHLYFPDDTLGLVDAENLRSLTCTENEL 504
QY 350 KFEQ---LCCHPFFSKIDMNNIRNSPPFVPTLKSDDDTSNF--DFPERNSVSSPQOL 404
Db 505 GRGGAHEIKSHAFRGVEFDLSLRIRAPFEPLTSAIDITPTPTDEIDQDNTALLKAQ 564
QY 405 SPGFGSG-----EE-----IPVGFYSVK 423
Db 565 AARGAAPAQAQEBSPSLSPFIFYGTPKR 592

RESULT 8

BEAL_HUMAN STANDARD; PRT; 1411 AA.
AC Q15075; Q14221;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Early endosome antigen 1 (Endosome-associated protein p162).
GN BEAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=95286647; PubMed=7769953;
RA Mu F.-T., Callaghan J.M., Steele-Mortimer O., Stenmark H.,
RA Parton R.G., Campbell P.L., McCluskey J., Yeo J.-P., Tock E.P.C.,
RA Toh B.-H.;
RT "BEAL, an early endosome-associated protein. BEAL is a conserved
RT alpha-helical peripheral membrane protein flanked by cysteine
RT 'fingers' and contains a calmodulin-binding IQ motif.";
RL J. Biol. Chem. 270:13503-13511(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Seelig H.P.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP INTERACTION WITH RAB5A.
RX MEDLINE=98361239; PubMed=9697774;
RA Simonsen A., Lippe R., Christoforidis S., Gaullier J.-M., Brech A.,
RA Callaghan J.M., Toh B.-H., Murphy C., Zerial M., Stenmark H.;
RL "BEAL links PI(3)K function to Rab5 regulation of endosome fusion.";
RN Nature 394:494-498(1998).
RN [4]
RP INTERACTION WITH RAB5A AND RAB5B.
RX MEDLINE=99421752; PubMed=10491193;
RA Callaghan J.M., Nixon S., Buccì C., Toh B.-H., Stenmark H.;
RT "Direct interaction of BEAL with Rab5b.";
RL Eur. J. Biochem. 265:361-366(1999).
RN [5]
RP INTERACTION WITH STX6, AND SUBCELLULAR LOCATION.
RX MEDLINE=99436077; PubMed=10506127;
RA Simonsen A., Gaullier J.-M., D'Arrigo A., Stenmark H.;
RT "The Rab5 effector BEAL interacts directly with syntaxin-6.";
RN J. Biol. Chem. 274:28857-28860(1999).
RN [6]
RP MUTAGENESIS OF ASP-1352; ASN-1357; L367-VAL-THR-1368; ARG-1375 AND
RP ARG-1400, HOMODIMERIZATION, AND INTERACTION WITH PHOSPHATIDYLINOSITOL
RP 3-PHOSPHATE.
RX MEDLINE=99322673; PubMed=10394369;
RA Kutateladze T.G., Ogburn K.D., Watson W.T., de Beer T., Emr S.D.,
RT Burd C.G., Overduin M.;
RL "Phosphatidylinositol 3-phosphate recognition by the FYVE domain.";
RN Mol. Cell 3:805-811(1999).
RN [7]
RP MUTAGENESIS OF TRP-1349; CYS-1358; PHE-1365; ARG-1370; ARG-1371;
RP HIS-1372; HIS-1373; CYS-1374; ARG-1375; CYS-1377; GLY-1378; CYS-1385;
RP ARG-1400 AND CYS-1405, SUBCELLULAR LOCATION, AND INTERACTION WITH
RP PHOSPHATIDYLINOSITOL 3-PHOSPHATE.
RX MEDLINE=20387352; PubMed=10807926;
RA Gaullier J.-M., Roennig E., Gillooly D.J., Stenmark H.;
RT "Interaction of the BEAL FIVE finger with phosphatidylinositol
RT 3-phosphate and early endosomes. Role of conserved residues.";
RN J. Biol. Chem. 275:24595-24600(2000).
RN [8]
RP INTERACTION WITH RAB22A.
RX MEDLINE=21859373; PubMed=11870209;
RA Kauppi M., Simonsen A., Bremnes B., Vieira A., Callaghan J.M.,
RA Stenmark H., Olkkonen V.M.;
RT "The small GTPase Rab22 interacts with BEAL and controls endosomal
RT membrane trafficking.";
RN J. Cell Sci. 115:899-911(2002).
RN [9]
RP MUTAGENESIS OF GLU-39; PHE-41; ILE-42; PRO-44; MET-47 AND TYR-60,
RP HOMODIMERIZATION, AND INTERACTION WITH RAB5C.
RX MEDLINE=22499593; PubMed=12493736;
RA Merithew E., Stone C., Bathiraj S., Lambright D.G.;
RT "Determinants of Rab5 interaction with the N terminus of early

FT	DOMAIN	335	1011	GLU-RICH.
FT	DOMAIN	1855	2107	GLU-RICH.
FT	DOMAIN	2178	2225	GRIP.
FT	VARIANT	61	61	T -> R (in strain Czech II).
FT	VARIANT	280	280	G -> S (in strain Czech II).
FT	VARIANT	293	293	G -> E (in strain Czech II).
FT	VARIANT	638	638	G -> S (in strain Czech II).
FT	VARIANT	819	819	K -> T (in strain Czech II).
FT	VARIANT	829	829	A -> T (in strain Czech II).
FT	VARIANT	859	859	T -> A (in strain Czech II).
FT	VARIANT	2065	2065	R -> S (in strain C57BL/6).
SEQ	SEQUENCE	2238 AA;	257562 MW;	494EA2C111F0165B CRC64;
Query Match				
Best Local Similarity 19.5%; Score 575.5; DB 1; Length 2238;				
Matches 419; Conservative 321; Mismatches 612; Indels 799; Gaps 90				
QY	4	FKYGARNFLDAGAAEPIASRASRLNL	-----FFGKPPPMFTQQOMSPLSREGILD	53
DB	98	FRSPIKESLFRSSKEPIVRTSRESLNQLDLDCAAAPDPSPDSESEADAPWSDGLSR	157	
QY	54	ALFVLVPECSOPALMKIKHNVFKVYSDTIAELQELQPSAKDFVRSVLVCGGHFAEVQV	113	
DB	158	-----EQILQRLRMERSLSYRGKYSILVTFQTILQ	-----	189
QY	114	VREKATGIYAMKMKKALLAQEQVFFEEERNILSRSTSPWIPOLQYAFODKXHLVLM	173	
DB	190	-REK-----KKUGILSSQSQ-----DKSL--RRISEUREBLQMDQAKXHL	---227	
QY	174	EYQPGDLLSLNARYEDQLDENLIQFYLAELILAVHSLMGVYHVDIKPENILVDRTG	233	
DB	228	-----QDEPDACL-----EEDQYISVLQT-247		
QY	234	HIKLYDFGSAAMKSNKMWNAKLPIGTPDYMAP-EVLTVNGDGKCTYGLDCDWSVGVI	292	
DB	248	-----QVSLKLRQLONGPMVADAPKLPPGELQAEVHGDKTEWEG-----VG--289		
QY	293	AYEMTYGRSPAEGTSARTFNINMFRFLKFPDDPKVSSDFDLIQ-----339		
DB	290	-----EPVGGTSAKTEML--QQRVKROE-----NLLQRCKETIGSHKQC	329	
QY	340	SLLCQKRLKFBGLCHCFFPSKIDWNINRSPPPVPTLKSDDDTSNDFDEPKNSWYSS	399	
DB	330	ALLLSKEALQ-----EQJD-----ERLQELERKMKELHMAEKLKIT-366		
QY	400	SPCQLSPGSGSEELPFVGFYSYKALGILGRSSVSYSGL-----DSPAKTSSMEKKLLTKS	455	
DB	367	---QLRDAKNLIEQL-----EODKGMVITTKQMLETLEKDETAQLRSHIKOMTTQG	418	
QY	456	KELODSQDK-----CHKQEQEMTRLHRRVSEVAVLSQKEVELKASQTOR	500	
DB	419	EBLRQEKESRAAFELEKALSTAQCTEDAQRMMKEMDE-----QMKAVE-RASEBER	472	
QY	501	SILLEODLATYITECSLSKRSLEQARMEVSDEDDKALQLLH-----540		
DB	473	URLQHELRSRVQEAASMAK-----KNSBEQVAALQKHAEELASKBOELSRRLEABER	525	
QY	541	DIREQSR-----KLQETKEEQYQOQVEEMRL-----MMNOLLEEDLVSAARRSDLY	585	
DB	526	ELQEQMRLALEKRSRSEYVLKLTQKEEQOESLASELELOKKAILTESENKJLQELQAEABY	585	
QY	586	ESEURESLAABE-FYKKA TECQH-----KLLKAKQDGKPEVCEY-----624		
DB	586	TRILETSTLEKLSQESKTSQSEHLAVHLBAEKXNKINKELTALABQHRTEVEGLQOQODS	645	
QY	625	---AKLEKINAEQQLKTQELQEKI--EKAA--KERA-----BRELEKLNREDS	666	
DB	646	LWTELOLSLSQOHAANVELREKYQOEKDALLESKSLTCAHILQDMNEKTLLEKLDKKQME	705	
QY	667	SEGRTKKLVEA-----ERRH-----682		
DB	706	LESVSSSELSELRARDQLABELSVLRGDADKMKQALAELEERQRRHHQREVS+SEQQEL	765	

Db 1809 ----CLPEELEBKCKSLIVSPMG-----EETGNNTGVK-----Q 1840

QY 1474 GNNKVPNNKRGQOGRXIVYLEGSKVLYDNEAREAQORVSEFELCLPDGVSIHCA 1533

Db 1841 NNASVVDVSQKTLQEKELTCQALEQRVKESLDLVRGAHRLEVEKL----- 1888

QY 1534 VGASELANTAKAEKARADAKLGNLSLLKLEGGDRDLDMNCTLFFSQVVLVGTGTEGLYALN 1593

Db 1889 -----TLKYEKSQ-----SSQEQMDGENK-----CVEVLEDR-----FEENSQSHE 1924

QY 1594 VLKNSLTHVPGGAVGTQIIVYIDLEKLMIAEERALCLVDVKKVKQSIAQ 1644

Db 1925 IOSNVGT-VDGLRS-----DLESKL--TGAER-----DQKLSKEVAR 1959

RESULT 10

CEP2_HUMAN

ID CEP2_HUMAN STANDARD; PRT; 2442 AA.

AC Q9BV73; O14812; O60588; Q9H450;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Centrosomal protein 2 (Centrosomal Nek2-associated protein 1) (C-NAP1)

DE Centrosome protein 250 (Centrosome associated protein CEP250).

GN CEP2 OR CNAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1] CEP2_HUMAN

RP SEQUENCE FROM N.A. (ISOFORM 1), AND AUTOIMMUNE DISEASE.

RC TISSUE=Cervical carcinoma;

RX MEDLINE=98165428; PubMed=9506584;

RA Mack G.J., Rees J., Sandblom O., Balczon R., Fritzler M.J.,

RA Rattner J.B.;

RT "Autoantibodies to a group of centrosomal proteins in human autoimmune

RT sera reactive with the centrosome";

RL Arthritis Rheum. 41:551-558(1998).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), INTERACTION WITH NEK2,

RP AND SUBCELLULAR LOCATION DURING THE CELL CYCLE.

RC TISSUE=Placenta;

RX MEDLINE=98311641; PubMed=9647649;

RA Fry A.M., Mayor T., Meraldi P., Stierhof V.-D., Tanaka K., Nigg E.A.;

RT "C-Nap1, a novel centrosomal coiled-coil protein and candidate

RT substrate of the cell cycle-regulated protein kinase Nek2.";

RL J. Cell Biol. 141:1563-1574(1998).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Clegg S., Cobley V.E., Collier R.E., Connor B.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA Lhvasialho M.H., Ievers M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie L.J., Mclay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,

RA Stuke C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;

RL "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 3).

RC TISSUE=Placenta;

RX MEDLINE=22388257; PubMed=12477932;

RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Alschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [5]

RP PHOSPHORYLATION, AND INTERACTION WITH NEK2 AND PPP1CA.

RX MEDLINE=21060765; PubMed=10880350;

RA Helps N.R., Luo X., Barker H.M., Cohen P.T.W.;

RT "NIMA-related kinase 2 (Nek2), a cell-cycle-regulated protein kinase

RT localized to centrosomes, is complexed to protein phosphatase 1.";

RL Biochem. J. 349:509-518(2000).

RN [6]

RP PHOSPHORYLATION DURING CELL CYCLE.

RX MEDLINE=22135747; PubMed=12140259;

RA Mayor T., Hacker U., Stierhof V.-D., Nigg E.A.;

RT "The mechanism regulating the dissociation of the centrosomal protein

RT C-Nap1 from mitotic spindle poles.";

RL J. Cell Sci. 115:3275-3284(2002).

CC -!- FUNCTION: Probably plays an important role in centrosome cohesion

CC during interphase.

CC -!- SUBUNIT: Monomer and homodimer (Probable). Forms a complex in

CC vitro with both NEK2 kinase and the PPP1CC catalytic subunit of

CC protein phosphatase 1 (Pp1).

CC -!- SUBCELLULAR LOCATION: Component of the core centrosome. In

CC interphase cells, it specifically associates with the proximal

CC ends of both mother and daughter centrioles. Associates with the

CC centrosome in interphase cells. In mitotic cells, it dissociates

CC from the mitotic spindle poles. At the end of cell division, it

CC reaccumulates at centrosomes.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=3;

CC Name=1;

CC IsoId=Q9BV73-1; Sequence=D:displayed;

CC Name=2;

CC IsoId=Q9BV73-2; Sequence=VSP_007372;

CC Name=3;

CC IsoId=Q9BV73-3; Sequence=VSP_007370, VSP_007371;

CC Note=No experimental confirmation available;

CC -!- TISSUE SPECIFICITY: Ubiquitously and weakly expressed.

CC -!- PTM: Differentially phosphorylated during cell cycle.

CC Phosphorylation may regulate association/dissociation from

CC centrosome. During M phase of mitosis, C-terminal part is

CC phosphorylated by NEK2, suggesting that it may trigger the

CC dissociation from the mitotic centrosome. It is dephosphorylated

CC in vitro by the Pp1 phosphatase.

CC -!- DISEASE: Antibodies against CEP2 are present in sera from patients

CC with autoimmune diseases that developed autoantibodies against

CC centrosomal proteins.

QY	1647 LPAQPD---	SPNIFEA 1660
Dd	2423 TSPGPVLLHPSPTTQA	2439
 RESULT 11 GOA4 HUMAN		
ID	GOA4_HUMAN STANDARD; PRT; 2230 AA.	
CD	Q13439; Q13270; Q13654; Q14436;	
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	
DE	Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256	
DE	kDa golgin) (Golgin-245) (72.1 protein).	
GN	GOLGA4.	
OS	Homo sapiens (Human)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.	
RA	MEDLINE=96215236; PubMed=8626529;	
RX	Exlich R.; Gleeson P.A.; Campbell P.; Dietzsch E.; Toh B.-H.;	
RT	"Molecular characterization of trans-Golgi p230: a human peripheral	
RT	membrane protein encoded by a gene on chromosome 6p12-22 contains	
RT	extensive coiled-coil alpha-helical domains and a granin motif.";	
RL	J. Biol. Chem. 271:8328-8337(1996).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RA	Seelig H.P.;	
RA	Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.	
RL	[3]	
RP	SEQUENCE OF 131-2230 FROM N.A.	
RC	TISSUE=Placenta;	
RC	MEDLINE=96125112; PubMed=8537393;	
RA	Fritzler M.J.; Lung C.-C.; Hamel J.C.; Griffith K.J.; Chan E.K.L.;	
RT	"Molecular characterization of golgin-245, a novel Golgi complex	
RT	protein containing a granin signature.";	
RL	J. Biol. Chem. 270:31262-31268(1995).	
RL	[4]	
RP	SEQUENCE OF 524-672 FROM N.A.	
RC	TISSUE=Gastric fundus;	
RA	Balague C.;	
RL	This (1994), Instituto municipal de investigation medica, Spain.	
-!	FUNCTION: May play a role in vesicular transport from the trans-	
CC	Golgi.	
-!	SUBCELLULAR LOCATION: Cytoplasmic, peripheral membrane protein	
CC	associated with the trans-Golgi network.	
-!	ALTERNATIVE PRODUCTS:	
CC	Event-Alternative splicing; Named isoforms=4;	
CC	Comment-Additional isoforms seem to exist;	
CC	Name=1;	
CC	IsoId=Q13439-1; Sequence=Displayed;	
CC	Name=2;	
CC	IsoId=Q13439-2; Sequence=VSP_004272, VSP_004273;	
CC	Name=3;	
CC	IsoId=Q13439-3; Sequence=VSP_004274;	
CC	Name=4;	
CC	IsoId=Q13439-4; Sequence=VSP_004275;	
-!	DOMAIN: Extended rod-like protein with coiled-coil domains.	
-!	DISEAS: Antibodies against GOLGA4 are present in sera from	
CC	patients with Sjogren's syndrome (SS) [MIM:270150]. Sera from	
CC	components of the Golgi complex.	
-!	DISEAS: Antibodies against GOLGA4 are found in sera from	
CC	hepatitis B patients.	
-!	SIMILARITY: Belongs to the golgin family.	
-!	SIMILARITY: Contains 1 GRIP domain.	
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way		

427	EBARRKKAEMDEQIKTIKTSBERISLQQLSRVQEVVDVVMKKSSEEQIAKLQKLHE	486	OC	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
692	ETMERRENKDDIQKTSQIQOMADKILE-----LEKKHREAVSAQHLEHV----	739	OX	NCBI_TaxID=44689;
487	KELARKEQLTKLQTRERFQEQMKVALEKQSEYLSIKSQEKEQESLAELELQKKA	546	RN	SEQUENCE FROM N.A.
740	-----LKQEQHVEEKIKVLNQIKDLADKETLENMQHREHEAEKXKILSE	788	RP	MEDLINE=87092266; PubMed=3540939;
547	ILTESNKLRLQQAETRYTRILESSLEKSLQENKQSKDLAVHLEA--EKXKNKE	604	RA	Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
789	QKAMINAMDSKIRSL-----QRIVELSEANK-----LAANSFLF	823	RT	"Conserved protein domains in a myosin heavy chain gene from
605	ITVMVEKHTELESUKHQODALWTEKLVQLKQYQTEMKLEKCEQKTELTKDKEIF	664	RL	Dictyostelium discoideum";
824	TORNKKAQEMISELSRQKXFLYETQAGLE-----AQRKLEBEQLEKSHQSDKXRLLEL	880	RN	Prcc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
665	QAHIEBNEKTEKLDVKQTELESSELSEVLKARHKLEELSVL--KQOTDKMK-QEL	721	RP	PHOSPHORYLATION SITES, AND MUTAGENESIS.
881	ETFLRVSLEHBEQ-----KLEKRLQTELSQERESQLTALQARAA	925	RC	STRAIN=AX2;
722	BAKDQKQKHQOQVDSIIKEHEVSVQTEKALKQOINQLELLKXERDKLKEHQAVEN	781	RX	MEDLINE=90353583; PubMed=2387408;
926	LESQLRQAETELETAEAE--BEIQALFAHREIQKFDALRNCTVITDLBEQNLQIT	983	RA	Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
782	LEADIKRSEGLQQAASAKLDVFSQSATHEQTKA-----YEEQLAQIQ	825	RA	Gerisch G.;
984	EDNAELNNQNFYLSKOLDEASGANDEI-----VQLRSEVDHLRREITEREMOLTS--	1033	RT	"Replacement of threonine residues by serine and alanine in a
826	QKLLDLETERILLTKQVAEVEAKQKDVCTELDAHKIQVDLMQOLEKQNSENEQKVKSLT	885	RL	phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
1034	-----QKOTMEAL-----XTTCTMLEEQVMDLEALNDBELLEKEROMEAWRSVL	1076	RN	FEBE Lett. 269:239-243(1990).
886	QVYESKLEDNKQEQTKQILVEKENMILQMRGQKKEITLTKLSAKEDSIH-----IL	941	RP	PHOSPHORYLATION SITES.
1077	GDF-KSQFCRVRELMDDTEKQSRADQRITE-----SRQVVELAVKSHK--A	1124	RX	MEDLINE=88112226; PubMed=2828113;
942	NEEYETKFNQEKMKVKQKAKEMQETLKKLLDQEAKLKKELENTALELSQKQFNA	1001	RA	Wagle G., Noegel A., Scheel J., Gerisch G.;
1125	EIALQALQALK-----EQKLKAEISLSP-----KLNDL-----EKK-----HAMLEM	1159	RT	"Phosphorylation of threonine residues on cloned fragments of the
1002	KMLEVAQANSAGISDAVSRLTNQKQIESLTVHRRLNDVSIWEKKLNQQAELQEI	1061	RL	Dictyostelium myosin heavy chain.";
1160	NARSLOQKLETERELKORLL-----EQAK-----LQOQMDLQKNH	1195	RN	FEBE Lett. 227:71-75(1988).
1062	HEIQLQKEQEAELKQKILLFGCEKEMNKETIWLKEGVKQDTTLNLEQLQKQSAH	1121	RP	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
1196	IFRLTQG--LOEALDR--ADLLKT--ERSDLEYQLENTQVLYSHEKVKVEGTISQOTKL	1248	RX	MEDLINE=95345066; PubMed=7619795;
1122	VNSLAQDFTKLKAHLEKLEVDLNSLKENTFLOEQLVLEKMLAEEDKRVSELTSLKTT	1181	RA	Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
1249	IDFLQAKMDQPAKKKGLFSRRKEDPALTPQVPLQVNLKALKEKAPCAELEBALQKT	1308	RT	Rayment I.;
1182	DEBFQSLKSHKSNKSLDKSLFKKLSELAIAQ--LDICCKKTEA-----LLEAKTNE	1234	RA	Smith C.A., Rayment I.;
1309	RIBLSAEBAARKATDHPHPT-----PATARO-----QIAMSALVRS	1348	RT	"X-ray structure of the magnesium(II).ADP.vanadate complex of the
1235	LINISSKTNAILSRISHCHORHTTKVKEALLIKTCVLSLELAQLRLTEQNTLNSIFQ	1294	RL	Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
1349	PEHQ-----PSAMSLAPPSSRRKE-----SSTPEFSRLKERMHNI	1387	RN	Biochemistry 35:5404-5417(1996).
1295	ATHQLEBKENQIKSMKADIESLVEKALQKEGNGQQAASEKSCITQLKELSEN	1352	RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RESULT 12			RX	MEDLINE=97452580; PubMed=9305951;
MYSD2 DICDI			RA	Guilick A.M., Bauer C.B., Thoden J.B., Rayment I.;
AC P08759;			RT	"X-ray structures of the MgADP, MgATPgammAS, and MgAMPPNP complexes
DT 01-NOV-1988 (Rel. 09, Created)			RL	of the Dictyostelium discoideum myosin motor domain.";
DT 01-OCT-1989 (Rel. 12, Last sequence update)			RN	Biochemistry 36:11619-11628(1997).
DT 15-MAR-2004 (Rel. 43, Last annotation update)			RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
DE Myosin II heavy chain, non muscle.			RX	MEDLINE=98070605; PubMed=9405148;
GN MHCA.			RA	Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
OS Dictyostelium discoideum (slime mold).			RT	"X-ray crystal structure and solution fluorescence characterization
			RL	of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
			RN	Dictyostelium discoideum myosin motor domain.";
			CC	J. Mol. Biol. 274:394-407(1997).
			CC	-!- FUNCTION: Myosin is a protein that binds to actin and has ATPase
			CC	activity that is activated by actin.
			CC	-!- SUBUNIT: Myosin II heavy chain is two-headed. It self-assembles
			CC	into filaments. Hexamer of 2 heavy chain subunits (MHC), 2 alkali
			CC	light chain subunits (MLC) and 2 regulatory light chain subunits
			CC	(MLC-2).
			CC	-!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
			CC	CORTEX.
			CC	-!- DOMAIN: Each myosin heavy chain can be split into 1 light
			CC	meromyosin (LMM) and 1 heavy meromyosin (HMM). It can be further

split into 2 globular subfragments (S1) and 1 rod-shaped subfragment (S2).

-!- DOMAIN: The rodlike tail sequence is highly repetitive, showing cycles of a 28-residue repeat pattern composed of 4 heptapeptides, characteristic for alpha-helical coiled coils.

-!- PTM: Phosphorylation inhibits thick filament formation and reduces the actin-activated ATPase activity.

-!- MISCELLANEOUS: Dictyostelium myosin II has no K(2)EDTA ATPase activity, perhaps correlated with the absence of a Cys at the SH-1 position (688).

-!- SIMILARITY: Contains 1 myosin-like globular head domain.

-!- SIMILARITY: Contains 1 IQ domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M14628; AAA33227.1; -;
 PIR; A26655; A26655.
 PDB; 1MWA; 03-DEC-97.
 PDB; 1MMD; 17-AUG-96.
 PDB; 1MMG; 03-DEC-97.
 PDB; 1MNN; 03-DEC-97.
 PDB; 1MND; 17-AUG-96.
 PDB; 1MNE; 17-AUG-96.
 PDB; 1VOM; 23-DEC-96.
 PDB; 1LVK; 28-JAN-98.
 PDB; 1DOX; 20-DEC-00.
 PDB; 1DOY; 20-DEC-00.
 PDB; 1DOZ; 20-DEC-00.
 PDB; 1D1A; 20-DEC-00.
 PDB; 1D1B; 20-DEC-00.
 PDB; 1D1C; 20-DEC-00.
 PDB; 1FMW; 20-DEC-00.
 PDB; 1G8X; 17-JAN-01.
 PDB; 1JWY; 07-NOV-01.
 PDB; 1JX2; 07-NOV-01.
 DictyBase; DDB0002015; mhca.
 InterPro; IPR000048; IQ region.
 InterPro; IPR001609; myosin head.
 InterPro; IPR004009; Myosin N.
 InterPro; IPR008989; Myosin_S1_N.
 Pfam; PF00612; IQ; 2.
 Pfam; PF00663; myosin_head; 1.
 Pfam; PF02736; Myosin_N; 1.
 PRINTS; PR00193; MYOSINHEAVY.
 ProDom; PD000355; myosin_head; 1.
 SMART; SM00015; IQ; 1.
 SMART; SM00242; MYSC; 1.
 PROSITE; PS50096; IQ; 1.
 KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
 KW Calmodulin-binding; Methylation; Phosphorylation.
 FT DOMAIN 1 761
 FT MYOSIN HEAD-LIKE.
 FT DOMAIN 762 791
 FT COILED COIL (POTENTIAL).
 FT NP BIND 817 2116
 FT ATP 186
 FT DOMAIN 638 660
 FT ACTIN-BINDING.
 FT DOMAIN 738 752
 FT ACTIN-BINDING.
 FT MOD_RES 130 130
 FT METHYLATION (DI-) (POTENTIAL).
 FT MOD_RES 1823 1823
 FT PHOSPHORYLATION (BY MECK).
 FT MOD_RES 1833 1833
 FT PHOSPHORYLATION (BY MECK).
 FT MOD_RES 2029 2029
 FT PHOSPHORYLATION (BY MECK).
 FT TURN 3 5
 FT TURN 7 8
 FT HELIX 10 15
 FT HELIX 25 28
 FT STRAND 34 37
 FT TURN 41 42

FT STRAND	48	55
FT STRAND	59	63
FT STRAND	69	73
FT HELIX	74	76
FT STRAND	78	79
FT HELIX	83	85
FT TURN	86	87
FT STRAND	90	90
FT HELIX	91	93
FT HELIX	99	110
FT TURN	111	113
FT STRAND	116	119
FT TURN	120	121
FT STRAND	122	126
FT HELIX	137	142
FT TURN	143	145
FT HELIX	148	150
FT HELIX	155	169
FT TURN	170	170
FT STRAND	173	179
FT TURN	181	182
FT HELIX	185	200
FT HELIX	210	226
FT STRAND	227	228
FT TURN	234	235
FT STRAND	236	237
FT STRAND	240	247
FT TURN	249	250
FT STRAND	253	261
FT HELIX	265	268
FT TURN	269	269
FT TURN	273	274
FT STRAND	278	278
FT HELIX	279	287
FT HELIX	290	296
FT TURN	297	297
FT HELIX	301	303
FT TURN	305	307
FT TURN	316	317
FT HELIX	320	334
FT TURN	335	335
FT HELIX	338	355
FT TURN	356	356
FT STRAND	360	360
FT STRAND	368	368
FT HELIX	373	382

Query Match 5.2%; Score 540.5; DB 1; Length 2116;
 Best Local Similarity 21.5%; Pred. No. 5.5e-10;
 Matches 326; Conservative 281; Mismatches 505; Indels 405; Gaps 68;

Qy	58	LFEECSQPALMKIKHVSNFVRKYSYSDTIAELQELQPSAKQPEVRSLVGCGHFA-----EVQ	112
Db	534	VFNATDNTLITKLH-SHFSKQN-----AKYEPRFKTEF-----GVTHYAGQVMYEIQ	582
Qy	113	VVEEKATGDIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPQLOYAFQDKHLYL	172
Db	583	DWLEK-----NKDPLQODLELCFKSDSNVVK-----LFNDPN---I	617
Qy	173	MEZYQPGDLLSLNRYEDQDENLIQFVLAELILAVSHVLMGYVHRDIKPEN-----	226
Db	618	ASRAKGANFITVAAQYKEQ-----LASLMATLETNN--PHFVRCIIPNNKQLPAK	666
Qy	227	-----ILVDR---TGHIKLVDFGSAAKNNKMNNAKLPIGTDPY-----MAPEVLTVNN	273
Db	667	LEDKVVDLQRCNGVLEGIRI--TRKGFNRIYA-----DFVKRYLLAPNVPDAE	717
Qy	274	GDGKGT-----YGLDCDWSVGVIAYEIMYGRSPFAEGTSARTFNNINFORFLKFPDD	327
Db	718	DSQKATDAVLKHLNIDPEQYRFGIT--KIFFRAGQLARIEEARE-----QRI-----	762
Qy	328	PKVSSDFDLIQSLCGQXERLKFEGLCCHPFFSKIDWNIR-----NSPPPPV-----	376

Db 119 DRELNLVKGSGFKVIVQRKXDTGEVYAMKVLKXKHIVEHNEVHTLSERNILQKINHP 178
 Qy 156 WIPOLQYAFQDKNHLVLMEEYQFGDLSLLNRYEDQDLENLIQYLAELIASHVHLM 215
 Db 179 FLVNLNYSFQTEKLYFILDYNGGELFVHLQK-DKXFTEDRVRYGAEIVLALHLHLS 237
 Qy 216 GYVHRDIPENILVDRTHGKILVDGSAKNSKNWAKVPIGTPDYMAPEVLTVMGD 275
 Db 238 GVIYEDLPENULLTNEGHICMTDFGLCKEGLLTFTDKGTFCGPEYLAPEVL---QGN 294
 Qy 276 GKGTYGLCDWMSGVIAVEMYIGRSPFAEGTSARTFNINMFQRLFPDPPDKVSSDFL 335
 Db 295 G---YKGQVDWWSFGSLLYEMLTGLPPFVNQDVQEMRYKIM-MEKLSPFH--FISPDAR 347
 Qy 336 DLIQSLLCQKPK-RLKFEGLC-CHFFFSKIDWNI--RNSPPFPVPTLKSDDDTSNFD-- 389
 Db 348 SLEQLLERDPKRLADPNLIKRRHFFRSIDWEQLFKQNPFPFPFVNVKGSADTSQIDFV 407
 Qy 390 -EPKNSVSSVSPCOLSPGSGEGELPFGVFSY 421
 Db 408 FDEAPSLTMAGECALNQ-----QQKDFEGFY 436

RESULT 14
 KAIB SCHPO
 ID KAIB SCHPO STANDARD; PRT; 607 AA.
 AC Q09898;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable serine/threonine-protein kinase C24B11.11c (BC 2.7.1.-).
 GN SPAC24B11.11c.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor A., Cronin A., Davis P., Feltwell J., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leacher S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potaashin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC HIGHLY SIMILAR TO YEAST DBF20 AND DBF2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EXBL: Z67757; CAA91776.1; --
 DR PIR; S62556; S62556.
 DR GeneDB Spombe; SPAC24B11.11c; --
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase_I.
 DR Pfam; PF00433; pkinase_C; 1.
 DR ProDom; EDD00001; Prot_kinase; 2.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 FT ATP-binding; Phosphorylation.
 FT DOMAIN 208 508 PROTEIN KINASE.
 FT NP_BIND 214 222 ATP (BY SIMILARITY).
 FT BINDING 237 237 ATP (BY SIMILARITY).
 FT ACT_SITE 331 331 BY SIMILARITY.
 FT MOD_RES 219 219 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 607 AA; 70491 MW; D9285263C7341FID CRC64;
 Query Match 5.1%; Score 540; DB 1; Length 607;
 Best Local Similarity 28.6%; Pred. No. 1.7e-10;
 Matches 142; Conservative 85; Mismatches 120; Gaps 15;
 Qy 27 LNLFFQCGKPFMTCCQMSPLSREGILDALFVLFEECSQPALMKIKHVSFVRKYSD--T 83
 Db 149 LDHYFEQLHLYTRKQPARLFEEQL-----KEPDSRR-----DELVKYNGRERV 194
 Qy 84 IAEQLQPSAKDFEVRSLVGCGHFAEVQVVRKATGDIYAMKVKKALLAEQVSFFE 143
 Db 195 YLKRRTTRISHGDFQTITVQGGYGVSWLARKRDTKEIVALKIMKSVLHKMDEIRHVL 254
 Qy 144 BERNILSRSTSPWIPQLOIYAFQDKNHLVLMEEYQFGDLSLLNRYEDQDLENLIQYLA 203
 Db 255 TERDILTANSEWLVRLLYAFQDTSNIYLAEMFVPGDFTLLSN-SGLVRDHHAKFYAT 313
 Qy 204 ELIIAVHSVHLMGVHRDIKPENILVDRTHGKILVDGSAKNSKNW-----252
 Db 314 EMFLAIDALHQLGYIHRDLKPFENFLVCASHIKLTFGLSSGIISKKIESMKIRLQEVN 373
 Qy 253 NAKLP-----IGTPDYMAPEVLTVMGDGKGTGVLDCDWM 287
 Db 374 NVVVPERSMTERRQVFRLLSQDPVVAHSVGVSGFDYMAPEVLRGEN-----YNHSVDY 427
 Qy 288 SVGVIAVEMYIGRSPFAEGTSARTFNINMFQRLFPDPPDKVSSDFL 340
 Db 428 SLGCMTECUSGPPFPFSGSVNWTWSNLKNWRCFORPHYDDPRDLFEFNRDWDANDFVCH 487
 Qy 341 LLCQKRL-KFEGLCCHFFFSKIDWNIENS-PPFPVPTLKSD-----DDTSN-----387
 Db 488 CITDPKDFCSLKQVMQHPYFSKIDWNVRTAYRPPFVPLNGLNSEIDAGYDFDTNENDMS 547
 Qy 388 -----FDEPKNSVSSVSPCOLSPGSGEGELPFGVFSYKALGILG 429
 Db 548 KYKEVHEKQAAIANMWNVTNFKPKRNA-----FIGFTFRHQKNSHP 587
 Qy 430 RSESVVSGLDSPAKTSS 446
 Db 588 TSSS-SALSSPLSAPS 602

RESULT 15
 MYHA RAT
 ID MYHA_RAT STANDARD; PRT; 1976 AA.

AC Q9JLT0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Myosin heavy chain, nonmuscle type B (Cellular myosin heavy chain, type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).
 GN MYH10
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=20483650; PubMed=11027611;
 RA Yam J.W.P., Chan K.W., Li N., Hsiao W.L.W.;
 RT "Molecular cloning and functional analysis of the promoter region of rat nonmuscle myosin heavy chain-B gene."
 RL Biochem. Biophys. Res. Commun. 276:1203-1209(2000).
 CC -I- FUNCTION: Cellular myosin appears to play a role in cytokinesis, cell shape, and specialized functions such as secretion and capping.
 CC -I- SUBUNIT: Myosin is a hexameric protein that consists of 2 heavy chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2 regulatory light chain subunits (MLC-2).
 CC -I- DOMAIN: The rodlike tail sequence is highly repetitive, showing cycles of a 28-residue repeat pattern composed of 4 heptapeptides, characteristic for alpha-helical coiled coils.
 CC -I- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -I- SIMILARITY: Contains 1 IQ domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF139055; AAF61445.1; -;
 CC HSP; P10587; 1BR2.
 CC InterPro; IPR000048; IQ_region.
 CC InterPro; IPR01609; myosin_head.
 CC InterPro; IPR004009; Myosin_N.
 CC InterPro; IPR02928; Myosin_tail.
 CC InterPro; IPR02017; Spectrin.
 CC Pfam; PF00612; IQ; 1.
 CC Pfam; PF00063; myosin_head; 1.
 CC Pfam; PF02756; Myosin_N; 1.
 CC Pfam; PF01576; Myosin_tail; 1.
 CC PRINTS; PD00193; MYOSINHEAVY.
 CC ProDom; PD000355; myosin_head; 1.
 CC SMART; SM00015; IQ; 1.
 CC SMART; SM00242; MYSC; 1.
 CC PROSITE; PS50056; IQ; 1.
 CC PROSITE; PS50056; IQ; 1.
 CC Myosin; ATP-binding; Calmodulin-binding; Actin-binding; Coiled coil; Multigene family.
 KW MYOSIN HEAD-LIKE.
 FT DOMAIN 1 785
 FT DOMAIN 786 815
 FT DOMAIN 845 1976
 FT NP_BIND 178 185
 FT SEQUENCE 1976 AA; 228963 MW; E32708BF9BF2B470 CRC64;
 Query Match 5.1%; Score 538; DB 1; Length 1976;
 Best Local Similarity 23.2%; Pred. No. 6.2e-10;
 Matches 257; Conservative 218; Mismatches 355; Indels 278; Gaps 47;
 QY 443 KTSMSEKLLIKSKELQSDQKCHKE---QEMTRLHRRVSEAVLS---QKEVELKA- 495
 DB 846 QVTRQEEELQAKDELLKYKQTKVGELEEMERKHQQLLEKNILAEQLQAEELFAE 905
 QY 496 SETQRSL---EODLATYITESSLSRLEQARMEVSQDDKALQL-LHDTREQRKLOE 551
 DB

Db 906 AEEMARLAACKQLEBEILHLESRVEG--BEERNQILQNEKKXWQAHIQDLEEQDB-EE 963
 QY 552 IKEEYQAQVEEMRLMNMVQLBED--LVSARRRDLSELSRSLAAEERKATECOHK 609
 Db 964 GARQKQLQKVTAFAKIKWEEVLLDQNSKFIEKKLMEDRIA-----ECSSQ 1014
 QY 610 LLKADQKPEVGEYAKLEKINAEQOLKIQELOKLEKAAK-----ERAPRELE- 658
 Db 1015 LAEEERAK-----NLAKIRNKQEVMSIDLEERLKEEKTQOELEKAKKLDGTTDL 1067
 QY 659 -----KLQNRDSSEG-----IRKKLVAEAE-- 679
 Db 1068 QDQIAELQAQVDELKVLTKKEBELQALARGDDETLHKNNALKVARELQAQIAELQEDF 1127
 QY 680 -----RHSLENKVKRL-ETMERRENRLKDDIOT-----KSQIQOMAD--KILELEE 724
 Db 1128 ESEKASRNKAEKQKRDLSSELEALKTELEDTLDTAAQBELTKRQEVAEALKALEDET 1187
 QY 725 KHREAQVS-----AQHLEVLXQEQHYBEKIKVLDNQIKKDLADKE---TLNMMQRH 775
 Db 1188 KNHEAQIDMRQRHATALEBELSQEQAKRFKANLEKKNKQGLETDNKLACAEVKVLQVVK 1247
 QY 776 EEEAHEKGIKLSSEKAMINAMDSKIRSEQRIVELSE-ANKLAANSSLFTQRMKAQEM 834
 Db 1248 ABSEHKRKKL-----DAQVQLHAKVSGDRLRVELAEKANKL-----QNELDNVSTL 1295
 QY 835 ISELROQKPYLETQAGKLEAQNRLKEOLEKISHQDHSNRLLELE---TRLREVSLEH 891
 Db 1296 LEEAKKGMKFAKDAAGLESQLODTQELLQETFRQNLSSRIRQLEEEKNSLQEQEEE 1355
 QY 892 BEQKLEIKRQTLQELSLQERESQ-----LTAQAAR-----AALESQRQ---AK 934
 Db 1356 BEARKNLEKQVLAQSLQADTKKVDGDDLTGTEGLEEAKKKLKDVEALSQRLEEKVLAY 1415
 QY 935 TELSETTAEEBEELQALTA--HRDEI-----QRKFD-ALRNSCTVITDLEEQNLQTL 983
 Db 1416 DKLEKTKNRLQBELDLDITVDLHQRIQVSNLEKQKQKFDOLLAEEKISARYAEERDRAE 1475
 QY 984 EDNALNNQNFYISKQDEASGANDEI-----VQLRSEV-----DHLRREITEREQMLT 1032
 Db 1476 AEAREKETKALSARALEEALAEKESFERQNKQLRADMEDLMSSKDDVGNVHELEKSKR 1535
 QY 1033 SQKQTEALKTCTMLBEQVMDLEALNDELLEKERQWEAWRSVL-----GDEKSQFECH 1086
 Db 1536 ALBQOQVEEMR---TQLEELDELEQATEDAKLRLEVNQMAKQAFERDLQTRDQNEKFR 1592
 QY 1087 -----VRELQRLMDTEKQSRARADQRTESRQVVELAVKEHKAELIA----- 1128
 Db 1593 LLLKQVRELEAELEDERKQALAA---VASKKKWEIDLKLEAQIEAANKARDEVIKQLR 1648
 QY 1129 -----LQALKEQKLEAESLSDKLNDELKXHAMLENNARSLOQKL-----ET 1170
 Db 1649 KLAQAKMDYQRELEEARASDEIFAQSEKSEKKLSLEAILQLEBELASSEPARHAEQ 1708
 QY 1171 ER-EL-----KQRLLEEQ-----AKLQOQMDLQKNHIFRITQGLQBALDEADL 1212
 Db 1709 ERDELADEIANSAGSKSALLDEKRRLEEARIAQLEELSEESQSNMELNDRFRKTLQVDT 1768
 QY 1213 LKTE-----RSDLEYQLE-----NTQVLSHEKVKQVNGTISQOTKLIDFLQA 1254
 Db 1769 LNTLEAASRAAKSDNARQQLERONKELKAKQLEAGAVSKFKFATISALEAKIQLEE 1828
 QY 1255 KMDQPA-----KXKGLF-----GRRKEDPALPTQVPLQVYNELKALEKEK 1295
 Db 1829 QLEGEAKERAANKLVPRTEKKLEIFMQVEDERRHAD-----QYKE---QMEKAN 1876
 QY 1296 ARCALEALQKTRIELRSAREEAAHRK 1323
 Db 1877 ARMKQLKRLQLEEA--EBEATRANASRRK 1902

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:12:55 ; Search time 27 Seconds
(without alignments)
3925.486 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10490

Sequence: 1 MLKFKYGARNPLDAGAAEPI.....QLNGEIROQVEKSVLRDYC 2053

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2430	23.2	497	4	US-09-804-471A-2
2	2430	23.2	497	4	US-10-238-709-2
3	2165.5	20.6	494	4	US-09-804-471A-4
4	2165.5	20.6	494	4	US-10-238-709-4
5	1277	12.2	257	4	US-09-916-204-2
6	1253	11.9	1388	2	US-08-685-576-1
7	1250.5	11.9	1354	3	US-08-685-871-2
8	1249.5	11.9	1388	4	US-09-976-594-296
9	1247.5	11.9	1388	2	US-08-685-576-4
10	1173	11.2	251	4	US-09-916-204-4
11	1170	11.2	251	4	US-09-916-204-5
12	1170	11.2	251	4	US-09-916-204-6
13	1043.5	9.9	900	2	US-08-630-822A-62
14	1043.5	9.9	900	2	US-09-005-063-62
15	1043.5	9.9	900	4	US-09-171-156A-21
16	1043.5	9.9	900	4	US-09-004-730A-21
17	1043.5	9.9	900	4	US-08-981-799A-21
18	847.5	8.1	420	3	US-08-685-871-58
19	843.5	8.0	582	2	US-08-422-699A-9
20	843.5	8.0	582	2	US-08-422-708B-9
21	830.5	7.9	420	3	US-08-685-871-59
22	799.5	7.6	638	2	US-08-422-699A-11
23	799.5	7.6	638	2	US-08-422-708B-11
24	776	7.4	555	1	US-08-484-044-6
25	706.5	6.7	479	4	US-08-442-100-13
26	706.5	6.7	479	4	US-08-939-106-13
27	706.5	6.7	479	4	US-09-442-102-13

28	691	6.6	526	4	US-09-442-100-12	Sequence 12, Appl
29	691	6.6	526	4	US-08-939-106-12	Sequence 12, Appl
30	691	6.6	526	4	US-09-442-102-12	Sequence 12, Appl
31	677.5	6.5	719	3	US-09-588-256-2	Sequence 2, Appl
32	673	6.4	404	2	US-08-860-150-3	Sequence 3, Appl
33	673	6.4	404	3	US-09-338-132-3	Sequence 3, Appl
34	668	6.4	464	2	US-08-878-989-4	Sequence 4, Appl
35	668	6.4	464	3	US-09-272-796-4	Sequence 4, Appl
36	665.5	6.3	465	2	US-08-878-989-18	Sequence 18, Appl
37	665.5	6.3	465	2	US-08-860-150-7	Sequence 7, Appl
38	665.5	6.3	465	3	US-09-338-132-7	Sequence 7, Appl
39	665.5	6.3	465	3	US-09-272-796-18	Sequence 18, Appl
40	661	6.3	500	4	US-09-442-100-14	Sequence 14, Appl
41	661	6.3	500	4	US-08-939-106-14	Sequence 14, Appl
42	661	6.3	500	4	US-09-442-102-14	Sequence 14, Appl
43	630	6.0	620	4	US-09-442-100-11	Sequence 11, Appl
44	630	6.0	620	4	US-08-939-106-11	Sequence 11, Appl
45	630	6.0	620	4	US-09-442-102-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-09-804-471A-2

; Sequence 2, Application US/09804471A

; Patent No. 6479269

; GENERAL INFORMATION:

; APPLICANT: WEBSTER, Marion et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CLO01164

; CURRENT APPLICATION NUMBER: US/09/804.471A

; CURRENT FILING DATE: 2001-03-13

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 497

; TYPE: PRI

; ORGANISM: Human

US-09-804-471A-2

Query Match		23.2%	Score 2430;	DB 4;	Length 497;
Best Local Similarity		99.1%	Pred. No. 8.8e-118;		
Matches		464;	Conservative	3;	Mismatches 1; Indels 0; Gaps 0;
Qy	1	MLKFKYGARNPLDAGAAEPIASRASRLNLFQKPPFTQOQMSPLSREGILDALFVLFE	60		
Db	1	MLKFKYGARNPLDAGAAEPIANRASRLNLFQKPPFTQOQMSPLSREGILDALFVLFE	60		
Qy	61	ECSPALMKIKHVSFVRKYSDDTIAEQLOPSAKDFEVRSLVCGGHFAEVQVVRKATG	120		
Db	61	ECSPALMKIKHVSFVRKYSDDTIAEQLOPSAKDFEVRSLVCGGHFAEVQVVRKATG	120		
Qy	121	DIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPQYAFQDKXNHLVMEYQFGG	180		
Db	121	DIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPQYAFQDKXNHLVMEYQFGG	180		
Qy	181	DLLSLLNRYEDQDENLIQFYLAELIILAVSHVLMGVVHRDIPKENTILVDTGHIKLVDF	240		
Db	181	DLLSLLNRYEDQDENLIQFYLAELIILAVSHVLMGVVHRDIPKENTILVDTGHIKLVDF	240		
Qy	241	GSAAMNSKNVNAKLPIGTDPYNAPEVLTVMNGDGKGTGYGLDCDMSVGVYAYEMIYGR	300		
Db	241	GSAAMNSKNVNAKLPIGTDPYNAPEVLTVMNGDGKGTGYGLDCDMSVGVYAYEMIYGR	300		
Qy	301	SPFAEGTSARTFNIMNFQFLKFPDPPKVSSDFLDLIQSLLCQKRLKPEGLCCHPFF	360		
Db	301	SPFAEGTSARTFNIMNFQFLKFPDPPKVSSDFLDLIQSLLCQKRLKPEGLCCHPFF	360		
Qy	361	SKIDNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSSPCOLSPSGFSEELPFYGF	420		


```
Db 361 SKIDWNIRNSPPFVPTLKSDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Qy 421 YSKALGILGRSESVVGLDSPAQTSSMEKLLIKSKELQSDQKCHM 468
Db 421 YSKALGILGRSESVVGLDSPAQTSSMEKLLIKSKELQSDQKCHV 468

RESULT 2
US-10-238-709-2
; Sequence 2, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Human
US-10-238-709-2

Query Match 23.2%; Score 2430; DB 4; Length 497;
Best Local Similarity 99.1%; Pred. No. 8.8e-118;
Matches 464; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLKFYKYGARNPLDAGAAEPIASRASRLNLPFGQKPPMTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFYKYGARNPLDAGAAEPIASRASRLNLPFGQKPPMTQQQMSPLSREGILDALFVLFE 60

Qy 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120
Db 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120

Qy 121 DIYAMKVMKKKALLAQEOVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180
Db 121 DIYAMKVMKKKALLAQEOVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180

Qy 181 DLLSLNRYEDQDENLIQFYLAELIIVAVSHVMGVVHRDIKPNILVDRTGHIKLVDF 240
Db 181 DLLSLNRYEDQDENLIQFYLAELIIVAVSHVMGVVHRDIKPNILVDRTGHIKLVDF 240

Qy 241 GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMGDGGKGYGLDCDWSVGVIAIYEMIYGR 300
Db 241 GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMGDGGKGYGLDCDWSVGVIAIYEMIYGR 300

Qy 301 SPFAEGTSARTFNINMFQRELKFPDPPKVSDFLDLIQSLLCQKRLKPEGLCCHPFF 360
Db 301 SPFAEGTSARTFNINMFQRELKFPDPPKVSDFLDLIQSLLCQKRLKPEGLCCHPFF 360

Qy 361 SKIDWNIRNSPPFVPTLKSDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Db 361 SKIDWNIRNSPPFVPTLKSDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420

Qy 421 YSKALGILGRSESVVGLDSPAQTSSMEKLLIKSKELQSDQKCHM 468
Db 421 YSKALGILGRSESVVGLDSPAQTSSMEKLLIKSKELQSDQKCHV 468

RESULT 3
US-09-804-471A-4
; Sequence 4, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
```

```
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Mus Musculus
US-09-804-471A-4

Query Match 20.6%; Score 2165.5; DB 4; Length 494;
Best Local Similarity 87.8%; Pred. No. 3.6e-104;
Matches 411; Conservative 26; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MLKFYKYGARNPLDAGAAEPIASRASRLNLPFGQKPPMTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFYKYGARNPEASASEPIASRASRLNLPFGKPLMTQQQMSALSREGMLDALFALFE 60

Qy 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120
Db 61 ECSOPALMKIKHVSFVQKYSDTIAELRELOPSARDFEVRSLVCGGHFAEVQVVRKATG 120

Qy 121 DIYAMKVMKKKALLAQEOVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180
Db 121 DIYAMKVMKKKALLAQEOVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180

Qy 181 DLLSLNRYEDQDENLIQFYLAELIIVAVSHVMGVVHRDIKPNILVDRTGHIKLVDF 240
Db 181 DFLSLNRYEDQDESMIQFYLAELIIVAVSHVMGVVHRDIKPNILVDRTGHIKLVDF 240

Qy 241 GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMGDGGKGYGLDCDWSVGVIAIYEMIYGR 300
Db 241 GSAAKMNSNK-VDAKLPIGTPDYMAPEVLTVMNEDRRGTYGLDCDWSVGVIAIYEMVYK 299

Qy 301 SPFAEGTSARTFNINMFQRELKFPDPPKVSDFLDLIQSLLCQKRLKPEGLCCHPFF 360
Db 301 TPFTEGTSARTFNINMFQRELKFPDPPKVSSELDLLQSLLCVQKRLKPEGLCCHPFF 359

Qy 361 SKIDWNIRNSPPFVPTLKSDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Db 361 ARTDWNIRNSPPFVPTLKSDDTSNFDPEKNSWAFILCVPAEPLAFSGEELPFVGF 419

Qy 421 YSKALGILGRSESVVGLDSPAQTSSMEKLLIKSKELQSDQKCHM 468
Db 421 YSKALGILGRSESVVGLDSPAQTSSMEKLLIKSKELQSDQKCHV 467

RESULT 4
US-10-238-709-4
; Sequence 4, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-238-709-4

Query Match 20.6%; Score 2165.5; DB 4; Length 494;
Best Local Similarity 87.8%; Pred. No. 3.6e-104;
Matches 411; Conservative 26; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MLKFYKYGARNPLDAGAAEPIASRASRLNLPFGQKPPMTQQQMSPLSREGILDALFVLFE 60
```

Db 1 MLKFKYGVRRPPASASEPTASASRLNLFQGGKPLMTQQQMSALSREGMLDALFALEE 60
QY 61 ECSQPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATG 120
Db 61 ECSQPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATG 120
QY 121 DIYAMKWKKALLAQEQVSFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEEYQPGG 180
Db 121 DIYAMKWKKALLAQEQVSFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEEYQPGG 180
QY 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240
Db 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240
QY 241 GSAAXXNSNMVNAKLPIGTPDYMAPEVLTVMGDGKTYGLDCDWSVGVIAEYMIYGR 300
Db 241 GSAAXXNSNMVNAKLPIGTPDYMAPEVLTVMGDGKTYGLDCDWSVGVIAEYMIYGR 300
QY 301 SPFAEGTSARTFNNIMNFORILKFPDDPKVSSDFLDLIQSLICGGKRLKFEGLCHPPFF 360
Db 301 SPFAEGTSARTFNNIMNFORILKFPDDPKVSSDFLDLIQSLICGGKRLKFEGLCHPPFF 360
QY 361 SKIDMNNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPCQLSPGSGEELFVGVGS 420
Db 361 SKIDMNNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPCQLSPGSGEELFVGVGS 420
QY 421 YSKALGILGRSSVSGSLSPAKTSMEKKLIKKELODSQDKCHKM 468
Db 421 YSKALGILGRSSVSGSLSPAKTSMEKKLIKKELODSQDKCHKM 468

RESULT 5
US-09-916-204-2
; Sequence 2, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THERSOF
; FILE REFERENCE: CLO01184CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Human
US-09-916-204-2

Query Match 12.2%; Score 1277; DB 4; Length 257;
Best Local Similarity 99.2%; Pred. No. 9.3e-59;
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLKFKYGAARNPLDAGAAEPIASASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFKYGAARNPLDAGAAEPIASASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFE 60
QY 61 ECSQPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATG 120
Db 61 ECSQPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATG 120
QY 121 DIYAMKWKKALLAQEQVSFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEEYQPGG 180
Db 121 DIYAMKWKKALLAQEQVSFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEEYQPGG 180
QY 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240
Db 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240
QY 241 GSAAXXNSNMVNAKLPIGTPDYMAPEVLTVMGDGKTYGLDCDWSVGVIAEYMIYGR 300

Db 241 GSAAXXNSNMVNAKLPIGTPDYMAPEVLTVMGDGKTYGLDCDWSVGVIAEYMIYGR 300
QY 301 SPFAEGTSARTFNNIMNFORILKFPDDPKVSSDFLDLIQSLICGGKRLKFEGLCHPPFF 360
Db 301 SPFAEGTSARTFNNIMNFORILKFPDDPKVSSDFLDLIQSLICGGKRLKFEGLCHPPFF 360
QY 361 SKIDMNNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPCQLSPGSGEELFVGVGS 420
Db 361 SKIDMNNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPCQLSPGSGEELFVGVGS 420
QY 421 YSKALGILGRSSVSGSLSPAKTSMEKKLIKKELODSQDKCHKM 468
Db 421 YSKALGILGRSSVSGSLSPAKTSMEKKLIKKELODSQDKCHKM 468

RESULT 6
US-08-685-576-1
; Sequence 1, Application US/08685576
; Patent No. 5906819
; GENERAL INFORMATION:
; APPLICANT: Kaibuchi, Kozi
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Nakano, Takeshi
; APPLICANT: Ito, Masaaki
; APPLICANT: Takahashi, No. 5906819uaki
; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,576
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-325129
; FILING DATE: 20-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-17150
; FILING DATE: 05-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-131206
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/843
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1388 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-685-576-1

Query Match 11.9%; Score 1253; DB 2; Length 1388;
Best Local Similarity 26.1%; Pred. No. 9.8e-57;
Matches 394; Conservative 291; Mismatches 507; Indels 320; Gaps 50;

QY 15 GAAEPIA--SRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFECSQPALMKIK 71
Db 12 GAPEAVSGDGAGASR-----QRKLEALIRDRSPINVESLDCGLNPLVLDLDFPALRKNK 66
QY 72 HVSNFVRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATGDIYAMKWKKK 131
Db 67 NIDNFLNRYEKIKYKIRGLQKMAEDYDVVKVIGRGAFGEVQLVRHKASQKVYAMKLLSKF 126
QY 132 ALLAQEQVSFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEEYQPGGDLISLLNRYED 191
Db 127 EMKRSDSAFFWEERDIMAFAFNSFWVQLFCAPQDDKYLHYMVEYMPGGDLVNLMSNYD- 185

Db 23 SEVNSCLLDGLDALVYDLDPPALRKKNKIDNIDNLSRYKDTINKIRDLRMKAEDYEVVVKVI 82
Qy 104 GCGHFAEVQVREKATGDIYAMKMKKALLAQEQVFFEEERNILSRSTSPWIPOLQYA 163
Db 83 GRGAFGEVQLVRHKSTKRVYAMKLLSKFEMIKESDSGAFFWEERDIFAPANSFWVQLFYA 142
Qy 164 FQKKNHLYMEEVQPGDGLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDK 223
Db 143 FQDRYLYMWEYMPGGDLVNLMSNYD--VPEKWARFYTAEVVLAIDAIHSGFTHRDYK 200
Qy 224 PENILVDRTHGHIKLVDFGSAAKNSKNKYNNAKLPICTPDYMAPEVLTVMNGDGKGTGYGLD 283
Db 201 PDNVLDDSGHLKADFGTCMKMKGVRCVDTAVGTPTYISPEVLKSQGGD--GYGRE 258
Qy 284 CDWWSGVVYAYEMIGRSPFABGTSARTNNIMNFORFLKFPDDPKVSSDFDLILQSLIC 343
Db 259 CDWWSGVVLYEMLVGDTPFYADSLVGTYSKMMHKNLSLTFDDNDNISKEAKNLICAFLT 318
Qy 344 GQKERLKFEG--LCCHPFPSKID--WNNIRNSPPFPVPTLSDDDTSNFDPEKNSWVS 398
Db 319 DREVLRGRNGVBEIKRHLFFKNDONAWETLRTVAVVPDLSDSDTSNFDDEEDKGE 378
Qy 399 SS--PCOLSPSGSGEELPVPVGSYSKALGILGRSSWSGLDSPAKTSSMEKXLIKSK 457
Db 379 ETFFPI---PKAFVGNQLPFVGTY-----YSNRRYLSSANTPNDNRSSN-----ADKS 423
Qy 458 LODS--ODKCKMEQEMTLRHRVSEVAVLQKEVELKASETOR---SLEQD-LATVIT 512
Db 424 LQESLQTIYKUEEQ-----HNEMQLKDEMEQCKRTSNIKLDMKIMBELD 468
Qy 513 ECSSLRSLQARMEYSQBDKALQLLHDIREQSRKLOEIKQEYQAOVE-EMRLMMNQL 571
Db 469 EECGNORNLST---VSQIEKEMLLQHRINEYQKAEQ--ENEXRNVENEVSTLKQOL 523
Qy 572 EEDLYSARRSDIYSESLRESLAAEFKFKATEQHKLLKAKDQCKPEVGEYAKLEKIN 631
Db 524 -BDLKVKVQNSQL-----ANE---KLSQLQKOLEEANDLRTTESDTAVRLKSH 568
Qy 632 AEQQLKIQLQKLEKAAERAELEKLQNEREDSSEGIRKKLVBAEERRHSLKNVKRL 691
Db 569 TEMSKSISQL--BSLNRELQER--NRILENSKSTQDYQLQAILAEARRDRG-----618
Qy 692 ETWERENLKDIDTK---SQOIQOQVADKILEEKEHREAOVSAQHLVELKQEQHY 747
Db 619 -----HDSMIGDLOARITSLQEEVKHLKHNLEKVEGERKEAQDMLNHS---KEK---666
Qy 748 EEKIVLDNQIKKLADKETLENMMORHEEAHE---KGKILSEKAMINAMDSKIRSL 803
Db 667 -----NLEIDLNYK--LKSLOQLEQVEYNEHKVTARLTDKXQISEAKSVAMCEM 716
Qy 804 EQRIVELSEANKLAANSLSFTORNMKAQEMIS--ELRQCKFYLETQAGKLEAQRKLEQ 862
Db 717 EKKLKEEREAREKAENRVV---QIEKQCSMLDVLKQSQCKLEHLTGKNE---RMEDE 768
Qy 863 LEKISHQDSDKNRILLETRLEVSLEHEBOK-----LEKQLTELQLS- 908
Db 769 VKNLTQLQESQENKLLQLNELKTOAFADNKLGLEKQMKQEBINTLLLEAKR--LLEFELAQ 827
Qy 909 -----LOERESOLTAQAAALAESOLROAKTELETTAAEAEETICALTAREDE 957
Db 828 LTKQYRGNEGQRELQDLQLEABQYFSTLYKTQVKELKEIEKQENRNLKQLEQNEKET 887
Qy 958 IQRKFDALRNSCTVITDLEQNLQNTENAEILNNQNLYSKQLDPSAGANDEIVQLRSEV 1017
Db 888 LATQLD-----LAETKAESQLA--RGLLEQYFELTQESKKAASRN-----927
Qy 1018 DHLREITEREMQLTSQKOTWEALATCTMTLBEQVMDLEALNDELLEKQERQWEARSVLG 1077
Db 928 ---RQETDKO-----HTVSRLEANSMTKDIILRRNEELTEKOKKAEYKLEK 977
Qy 1078 DE-----KSQFECRVRELQRLDTEKQSRARADORITESQVVELAVKHAELALQQA 1132
Db 978 EEEISNLKAAFE-----KNINTERTLKTQAVNKLAEIMNRKOPKIDRKKANTQDLRKK 1030

Qy 1133 LKEQK---LKAESLSDKLNLDLEKKHAMLEMNARSIOQKLETERELKORLLEBEQA---KLQ 1186
Db 1031 EKENKQLQELNQEREXFNQVVKH-----QKEL---NDMQAQLVEECAHRELQ 1077
Qy 1187 QOMDLQKNHIFRLTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVMKEGTISQOT 1246
Db 1078 MCLASKESDIBQLRAKLDLSDTSVASFPFAD-----ETDGNLPESR 1120
Qy 1247 KLIDFLOAKMDQPAKKKKGLFSRRKEDPALPTQVPOVLYNELKLALEKEKARCALEALQ 1306
Db 1121 -----IEGWLSPVNRGNIKRYGKWKQVYVSSKILFYNDQEKQNPNSVLDIKL-- 1173
Qy 1307 KTRIELSAREEAHRKATDHPHPSTPATARQQIAMSALVRSRPHQPSAVSLLAPPSRR 1366
Db 1174 ---FHVPRVTOGDVYRAET-----EEIFKIFQILYANEDEC 1206
Qy 1367 KESSTPEEFSRLKERVHHNIPHFVNLNMRATKCAVCLDTVHFGQASKCLEQVWCH 1426
Db 1207 RQVEMEPVQOAKETNFQNHKGHEFIPTLYHFPANCDACAKPLMHVFKPPPALECR-RCH 1265
Qy 1427 PKCS-----TCLPATCGLPAEYATHFTFAFCRDKMNSPGLQTKPSPSSLHLEGMM 1476
Db 1266 VKCHRDHLDKKEDLIC-----PCKVSYDVTSA--RDMLLACSDQDEQKQVWTHLV--K 1314
Qy 1477 KVPNNKRG 1485
Db 1315 KIPKNPPSG 1323

RESULT 8

US-09-976-594-296
; Sequence 296, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 296
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1
US-09-976-594-296

Query Match 11.9%; Score 1249.5; DB 4; Length 1388;
Best Local Similarity 26.3%; Pred. No. 1.5e-56;
Matches 391; Conservative 291; Mismatches 513; Indels 293; Gaps 47;

Qy 24 ASRLNLFQCKPFPMTQQQSPLSRSGILDALFVLEECSPALMKIKHVSFVRYSDT 83
Db 24 ASR-----QKLEALIRDPSPINVSLLDGLNSLVLDLDFPALRKNKNDINFLRYEKI 78
Qy 84 IAELOELQPSAKDFEVRSLVCGGHFAEVQVVRKATGDIYAMKMKKALLAQEQVSFFE 143
Db 79 VKKIRG-QMKAEDYDVVKVIGRGAFGEVQLVRHKAQKQVYAMKLLSKFEMIKESDSAFFW 138
Qy 144 EERNILSRSTSPWIPOLQYAFQDKNHYLMEEYQPGDGLLSLNRYEDQDENLQFYLA 203
Db 139 EEDINAFANSPWVQLFYAFQDDRYLYMMEYMPGGDLVNLMSNYD--VPEKWARFYTA 196
Qy 204 ELILAVHSVLMGYVHRDKPENILVDRTHGHIKLVDFGSAAKNSKNKYNNAKLPICGTDPY 263
Db 197 EVVLALDAHSMGLIHRDVKPDNKLIDKGLKLADFGTCMKQDETCMWHCDTAVGTDPY 256

QY 264 MAFVLTVMGDKGTGYGLDCDWMVGVYAYEMIVGRSPFAGTSGARTENNTMNFQRLK 323
DB 257 ISPEVLKSGQGD--GYFGRECDWMVGVFLYEMLVGDTFTFYADSLVGTYSKIMDHNSLC 314
QY 324 FPDPPKVSDFDLIOSLCQKXERLKFEG--LCCHPEF--SKIDMNNIRNSPPFFVPT 378
DB 315 FPDALISKHAKNLI CAFLTDREVLGRNGVEIRQHPFFKNDQWHDNIRETAAPVVP 374
QY 379 LKSDDDTSNPDPEKNSW--VSSPCQLSPGSGEELPFVGSYSKALGILGRSSEVSG 437
DB 375 LSSDIDSSNPDIDDKGVETPPI---PKAFVGNQLPFIFGTYYRENLLLS----- 423
QY 438 LDSPA--KTSSMEKLLIKSQDQKCKKQEQMTRLHRRVSEAVLSQKEVELKA 495
DB 424 -DSPSCRENDISQSR--KNEESQETQKLYTLBEHLSN-----EMQA 462
QY 496 SETQSLLEODLATYITECSSLRSLEQARMEVSQE-----DDKALQLLHDIRQSRKL 549
DB 463 KEE-----LEQ-----KCKSVNTRLEKTAKELEEBITLRKSVESALRQ-----ERKAL 507
QY 550 QETKEOEYQAEEMRLMWNQLEEDLVASARRRSDLYESELRESRLAAEFERKATECOHK 609
DB 508 LOHKNABYQKADHEADKGNLNDVNSLKQLEDLKKKNQNSQISTE-----KYNQLORQ 563
QY 610 LLKAKQDGKPEVGEYAKLEKINAEQOLKIQ-----ELQEK--LEKAAKERAERELEKL 660
DB 564 LDETALLRTESDTAARLRKTAQESSKQIQLESNNRDLQKNCLLETAKLKEFEFINL 623
QY 661 QNRDSESGIRKLVAEERRHSLKVNKLEMTERRNRLKDDIQTKSQOIQQWADKIL 720
DB 624 QSALESE---RRDTHGSEIINDLQRIQGLEEDLKNKILLAKVELEKRRQLEPTD-L 679
QY 721 ELBEKREAOVSAQ--HLEVLHKKQE-QHYEEIKVLD-NOIKKOL--ADKETLENMQR 774
DB 680 EKEKSNWEIDMTYQLKVIQSLQSEBAEHKATKARLADKNKIVESIEEAKSEAMKEMKK 739
QY 775 HEBEAHEKGI---LSEKAMINAMDSKIRSLRQIVELSEANKLAANSLSLTQRMKQA 831
DB 740 LLEERTLKQVENLLLEAEKRCSLPCDLKQSQOKINEL-----LKOK 782
QY 832 EEMISLRQOKFYLETOAGLEQNRKLEBQLEKISHQDHSKRNRLLELETRILREVSLEH 891
DB 783 DVLNEDVRNLTLEKQETQKCCITQNDLKNQTOVNTLKMSEK-QLKQENHLMEMKXNL 841
QY 892 BEQKLELRQUTLQLSQRESQLTALQARALESQRLQAKTELEETT---ABAESEI 948
DB 842 EKQNAELRKERQADGQMKELQLEAEQVFTLYKTQVRELKECEERTKLGLKELQKK 901
QY 949 QALTAHREDETRFDALNSCTVITDLBEQNLQITEDNALNNQNFVLSKQLEASGAND 1008
DB 902 QELQDRDSLAQLE-----ITLTKADSEQLARSIAEBEQVSDLEKEKIMKE----- 947
QY 1009 EIVQLRSEVDHLREITEREMOLTSQKTMEALKTTCTMLEEQVMDLEALNDELLEKQ 1068
DB 948 --LEIKEMVARKQELTEKDATIASLEETNRTLTSQVANLAN--EKEELNNKLKDVQSQ 1002
QY 1069 WEAWRSVLGE-----KQFECVRELQRLMDTEKQSPARADQRTESRQVVELAVKE 1121
DB 1003 ----LSRLKDEEISAAAKAQFE-----KQLLTERTLTKTQAVNKL----- 1038
QY 1122 HKABILAQALKEQKLKABSLDKLNDLEKHAMLEMMARSQOQKLETERELKQRLLEE 1181
DB 1039 --AEMNKEKPVK-----RGNDTDVRKK--EKENRKLHMLKSERE---KLTQ 1080
QY 1182 QAKLQOQMDLQKNHIFLTO---GLQALDRALLKTERSQDLEYQIENIQVLY----- 1231
DB 1081 MIXYQKELNEMQAIAESQIRLEQMTLDSK-----SDIEQLRSQALHIGLDS 1134
QY 1232 -----SHEKVMGEGTIS-----QOTKLIDFLQAKMDQPAKKKGLFSRRKEDPA 1275
DB 1135 IGSFGDAEADDGFPESLELGSWLSLVPVNTTKFGWKVYIVSSKILFYDSEQKEQS 1194

QY 1276 LPTQV-----PLQYNELKLALEKEKARCACAELEELAQKTRIELRSAREBAHRKAT 1325
DB 1195 NPYMVLIDIKLFHVRPVTQTDVYRADAKETPRIFQI----- 1230
QY 1326 DHPFSPATARQOIAMSAIVRSPHQPSAMSLAPPSSRRKESSTP-----EEFSRLKE 1381
DB 1231 -----JYANEGSSKKQEOPFVFPVGEKSNYICH 1258
QY 1382 RMHENIHRNFVGNMNAATKCAVCLDTVHFGRQASKCLEQVNMCHPKC 1429
DB 1259 KGHEFIFTLHYF-----PTNCEACMKPLMHMFPPPALECR-RCHIKC 1300

RESULT 9
US-08-685-576-4
; Sequence 4, Application US/08685576
; Patent No. 5906819
; GENERAL INFORMATION:
; APPLICANT: Kaibuchi, Kozo
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Nakano, Takeshi
; APPLICANT: Ito, Masaaki
; APPLICANT: Takahashi, No. 5906819uaki
; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,576
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-325129
; FILING DATE: 20-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-17150
; FILING DATE: 05-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-131206
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/843
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1388 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-685-576-4

Query Match 11.9%; Score 1247.5; DB 2; Length 1388;
Best Local Similarity 26.3%; Pred No 1.9e-56;
Matches 391; Conservative 290; Mismatches 54; Indels 293; Gaps 47;
QY 24 ASRLNLFQGGPPMTQOQMSPLSREGILDALFVLFECQOPALMKJHVSNFVKYSIDT 83
DB 24 ASR-----QRKLEALIRDPSPINVESILLDGLNSLVLDLDFPALRKNKNIDNFLNRYEKI 78

QY	84	IABELOQPSAKDFEVRSLVGVCGHFAEVQVREKATGDIYAMKVMKKKALLAQEQSVFEE	141
Db	79	VKKIKGLQMKAEEDYDVVKVIGRGAFEVQLVRHKASQKYAMKLLSKFENIKRSUSAFW	138
QY	144	BERNILSRSTSPWIPOLQYAFQDNKHLYIMEEYPOGDDLSSLNRYVEDOLDENLIQFYLA	203
Db	139	BERDINAFANSPWVQGLFYAFQDRLYLVMYMEYMGDLVNLMSNYD--VPEKWAIFYTA	196
QY	204	EUILAVHSVHLMGYVHRDILKPENILVDRTGHIKLVDFGSAAKNNSKMYNAKLPIGTDPY	263
Db	197	EVVLALDAIHSNGLIHRDVKFDNMULDKHGLKLADFGTCKMOWDETGMVCHCTAVGTDPY	256
QY	264	MAPEVLTVNMGOGKGTGYGLDCDWSVGVYIAYEMIIYGRSPFAEGTSARTFNNIMNFQFLK	323
Db	257	ISPEVLKSQGGD--GFYGRCDWSVGVFLYEMLVGDTFPYADSLVGTYSKIMDHKNSLC	314
QY	324	FPDDPKVSDFDLILQSLICGOKERLKFG--LCHPEF--SKIDMNNIRNSPPFPVPT	378
Db	315	FEDAEISKHAKNLCAFLTDREVLGRNGVEIRQHPFFQDQWHDNIRTAAPVPE	374
QY	379	LKSDDDTNFDPPEKNSW--VSSSPCOLGFSGSGEELPFVGFYSKALGILGRSESVSG	437
Db	375	LSSDIDSNNFDDIEDDKGVETFTPI--PKAFVGNLPFGTGYTYRENLLS-----	423
QY	438	LDSPA--KTSSMEKKLLIKSKLEQDSQDKCHWCEQEMTLHRVSEVAVLSQKEVELKA	495
Db	424	-DPSPRENDSTQSR--KNEESQIQTKVLYLEEHLN-----EMQA	462
QY	496	SETQSRLLQEODLATYITECSSLRKSRLEQARMEVSQZ-----DDRALQDLDHIREQSKL	549
Db	463	KBE-----LEQ-----KCKSVNTRLEKTAKELEEEITLRSKSVESALQQL-----ERKAL	507
QY	550	QEIKEQEQYCAQVEENRMLMNQLEEDLVARRSDLYESELRESRUAAEFKPKATECOHK	609
Db	508	LQHKNAEYORKADHEADKKNLENVDNLSKQOEDLKKRNQNSQISTE-----KVNQLQRQ	563
QY	610	LKAKDOQKPEYGEYAKLEKNAEQOLKIQ-----ELQEK--LEKAAXERAREBEKL	660
Db	564	LDEINALLRTESDTAARLRKTAQESSKIQOLESNNRDLQDNCLLETAKLKEFNL	623
QY	661	QNRDSSGEGIRKLVAEERHSLNKVKRLTETMERRENLRKDDITQTSQQIQQVADKIL	720
Db	624	QSALESE---RRDRTHGSEIINDQGRICGLEEDLNGKIKLLAKVELEKROLQRFRFTD-L	679
QY	721	ELEEKHRAQVSAQ--HLEVLHKOKE-OHYEEKIKVLD-NOIKDLD--ADKETLENMQR	774
Db	680	EKESXNVEIDMTYQLVKIQSLEQEEAEHKATKARLADKNKIYSIEFAKSEAMKEKK	739
QY	775	HEEZAHEKGKI---LSQKAVINAMWSKIRLSQIRVELSPANKLAANSLFTORNKQAQ	831
Db	740	LLEERTLKQKVENILLEAEKRCSLDCLDKSQSQKINEL-----LKQK	782
QY	832	EEMSELROQKPYLETOAGLEAQNKRLEQOLEKISHODHDKNRLLLETLRLSEVSLH	891
Db	783	DVLNEDVRLNLTKEQETQKCEKLTQNDLKMTQQVNTLKQSEK-OLKQENHLMEMKNL	841
QY	892	EEQKLELKRQUTELQLSQIQRESQITALQAARAALESQLRQAKTELEBTT---AEABEEI	948
Db	842	EKQNAELRKQDADGQMKELQDQLEABQYFSTLYKTQVRELKECEKTKLQKELQOKK	901
QY	949	QALTAHDEIQKPKDALRNSCTVITDLEEQNLQITEDNAELNNQNFYLSKQLDSEASAND	1008
Db	902	QELQERDLSLAQLE-----ITLTKADSEQLARSTAEQYSDLREKIMTK-----	947
QY	1009	EIVQLRSEVDHLRREITEREMQLTSQRTQMEALTKTCTMLEEQVMDLEALNDELLEKERQ	1066
Db	948	--LEIKENMARHKQELTEKDATIASLEETNRTLSDVANLAN---EXEELNNLKDQVQEQ	1002
QY	1069	WEAKRSVLGE-----KSOFECEVRELQRMILDEKQSRARADQRTESQVVELAVKE	1121
Db	1003	-----LSRLKDEEISAAAKQAFE-----KQLTERTLTQVANKL-----	1038

```

QY 1122 HKAETLALQQALKKQKLAESLSKDKNLDEKXHAMLEWNAARSLOOKLETRELRKQRLLEE 1181
DB 1039 --AELMNRKPEVK-----RGNDTDDRKK--EKENRKLHMLKESPRE--KLTTQ 1080
QY 1182 QAKLQQQMDLQKNHIFRLTQ--GLOALDRADLLKTERSDLEYQLENIQVLY----- 1231
DB 1081 MIKYQKELNEMOAOIABESQIRBELQMTLDSKD-----SDIEQLRSQALHIGLDS 1134
QY 1232 -----SHEKVKMEGTIS-----QOTKLIDLOAKMDQPAKKKGLFSRRKEDPA 1275
DB 1135 IGSFGDAEADGGPESRLEGWLSLFRVNTTKFGWVKVTVVSSKILFYDSEQDKEQS 1194
QY 1276 LPTQV-----PQYNELKALKEKARCABELEALQKTRIELRSAREEAAHRKAT 1325
DB 1195 NPYVNLIDIKLFHVRPVTQTDVVRADAKEIPRIFOI----- 1230
QY 1326 DHPHPSTPATARQOIAMSAIVRSPHQPSAMSLAPPSSRRKESSTP-----EFSRRLKE 1381
DB 1231 -----LYANEGESKKEQFEPVBPVGKSNYICH 1258
QY 1382 RMXHNHPRFNVGLNMRATKCAVCLDTVHFGROASKLEQVNVCHPKC 1429
DB 1259 KGHEFIPTLYHF-----PTNCEACMKFLWHMFKPPFALECR-RCHIKC 1300

RESULT 10
US-09-916-204-4
; Sequence 4, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-916-204-4

Query Match 11.2%; Score 1173; DB 4; Length 251;
Best Local Similarity 90.0%; Pred. No. 2e-53;
Matches 226; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLKFYTGARNPLDAGAAEPTIAGRASRLNLFQGGKPPMTQQQWSPLSRREGILDALFVLE 60
DB 1 MLKFYGVNPNPSEASAPEPIAGRASRLNLFQGGKPLMTQQQWSPLSRREGILDALFVLE 60
QY 61 ECSQPALMKIKHVSNEFKVYSDTIAELOLPSAKDFVRSILVCGGHFAEVQVREKATG 120
DB 61 ECSQPALMKIKHVSNEFKVYSDTIAELRELQPSVDFVRSILVCGGHFAEVQVREKATG 120
QY 121 DIYAMKVKKKALLAQEQVSPFEERNILSRSTSPWIPOLQYAFQDKNHLIYMEYQPGG 180
DB 121 DVIYAMKIMKAALRAEQVSPFEERNILSQSTSPWIPOLQYAFQDKNHLIYMEYQPGG 180
QY 181 DILSLNRYEDOLDENLIOFYLAELILAVHSVHLNGYVHROTKPENILVDRGHIKLVD 240
DB 181 DILSLNRYEDQDENNIQFYLAELILAVHSVHQGYVHRDTPENILIDRTGHIKLVD 240
QY 241 GSAAKGNVSNKM 251
DB 241 GSAAKGNVSNKV 251

```

RESULT 11
US-09-916-204-5
; Sequence 5, Application US/09916204

```

; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-916-204-5

Query Match      11.2%; Score 1170; DB 4; Length 251;
Best Local Similarity 89.6%; Pred. No. 2.9e-53;
Matches 225; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPTASRASRLNLFPGKPPFMTQQQMSPLSREGILDALFVLFE 60
DB 1 MLKFKYGVNRNPPASASEPTASRASRLNLFPGKPLMTQQQMSALSREGMLDALFALFE 60

QY 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120
DB 61 ECSQPALMKMKHVSFVQKYSYDTIAELRELQPSARDFEVRSLVGCCHFAEVQVVRKATG 120

QY 121 DIYAMKMKKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGG 180
DB 121 DVYAMKIMKKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGG 180

QY 181 DLSLLNRYEDQDENLQFYLAELILAVHSHVLMGYVHRDIKPENILVDRTHGHIKLVDF 240
DB 181 DFLSLLNRYEDQDENLQFYLAELILAVHSHVLMGYVHRDIKPENILVDRTHGHIKLVDF 240

QY 241 GSAAKMNSNKM 251
DB 241 GSAAKMNSKV 251

RESULT 12
US-09-916-204-5
; Sequence 6, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-916-204-6

Query Match      11.2%; Score 1170; DB 4; Length 251;
Best Local Similarity 89.6%; Pred. No. 2.9e-53;
Matches 225; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPTASRASRLNLFPGKPPFMTQQQMSPLSREGILDALFVLFE 60
DB 1 MLKFKYGVNRNPPASASEPTASRASRLNLFPGKPLMTQQQMSALSREGMLDALFALFE 60

QY 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120
DB 61 ECSQPALMKMKHVSFVQKYSYDTIAELRELQPSARDFEVRSLVGCCHFAEVQVVRKATG 120

QY 121 DIYAMKMKKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGG 180
DB 121 DVYAMKIMKKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGG 180

QY 181 DLSLLNRYEDQDENLQFYLAELILAVHSHVLMGYVHRDIKPENILVDRTHGHIKLVDF 240
DB 181 DFLSLLNRYEDQDENLQFYLAELILAVHSHVLMGYVHRDIKPENILVDRTHGHIKLVDF 240

QY 241 GSAAKMNSNKM 251
DB 241 GSAAKMNSKV 251

RESULT 13
US-08-630-822A-62
; Sequence 62, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, SHIRLEY WU
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDIA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,822A
; FILING DATE: 11-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CONNELL, GARY J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-630-822A-62

Query Match      9.9%; Score 1043.5; DB 2; Length 900;
Best Local Similarity 27.0%; Pred. No. 3.7e-46;
Matches 279; Conservative 221; Mismatches 364; Indels 171; Gaps 31;

QY 70 IKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATGDIYAMKMK 129
DB 1 MKSEAYTNREITASELVNLRMKPDDENLKVIRGAFGEVQLVRHKSTAQVFMKRLS 60

QY 130 KXALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGDLSLNRY 189
DB 61 KPEMIKRPDSAFFWEERHIMAHAKSEWIVQLFFAFQDKXYLYMVMYMPGGDLVSLMSDY 120

QY 190 EDQDENLQFYLAELILAVHSHVLMGYVHRDIKPENILVDRTHGHIKLVDFGSAKNSN 249
DB 121 E--IPEKWMVFTTMEVVLALDTIISMGFVHRDVKPDNMLDKYGHKLADFGTCMKMDTD 178

QY 250 KMVNAKPIGTPDYMAPEVLTVMGDNGKGTGVLGDCDMMWSGVIAVEMYIGKSPFAEGTSA 309

```


Db 179 GLVRSNNAVGTDPYISPEVL--QSOGGEVYGRECDWNSVGIPLYEMLFGETPFYADSLV 236
QY 310 RTFNINMNFQFLKFPDDPKVSSDFLDLIQSLLCGQKERL---KPEGLCCHPFFSKIDW- 365
Db 237 GIYSKIMDHNSITPPEVEISQYARSLQGLFDTQELGRNEVEEIKSHFFPINDQWT 296
QY 366 -NNIRNSPPFPVTLKSDDDTGNFDE-----PEKNSWVSSPCQLSPGSGFSGBELPFVG 418
Db 297 FDNLRDSAPPVPELSGDDDTNFDIERDETPEENFPI-----PKTFAGNHLPPVG 348
QY 419 FYSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKE-----LQDSQDKCHKEQ 470
Db 349 FTYNGDYQLLTN-----GGVRSNDMVDTKLNNICVSSKDDVNLQNLLEQKNGSENKLT 403
QY 471 EMTRLHRRVSEVAVLSQKEVELKASQETORSLEQDLATVITCSSLKRSLE---QARME 527
Db 404 NTQLLSNKLDE---LGQRECELR---NQAGDYKELTKFKLSCKELQKAEFENELRK 456
QY 528 VSQ---EDDKALQLLHDIREQSKLOEI-KEQBYQAOVBEEMLMNQLBEDLVSARRSD 583
Db 457 TESLLVETKRLD-----BEQNKRTREMNNOQHNDKINMLEKQINDLQEKL----- 503
QY 584 LYESELRESLAAEFKPKATECOHKLKAKDQKPEVGEYAKLEKINAEQOLKIQLQEQ 643
Db 504 --KGELEHN---QXLLKQCAVELR-----VAQSA-TEQLANLEQETWQGLQT 543
QY 644 KLEAKAKERAERELEKILQNRDSSGIRKLVAEERHSLENKVKRLTWERRENRLKD 703
Db 544 Q-----RDALQEVASLOKLSQERSRS--QASDMQIELEAKLOALHIELEHVRNCD 595
QY 704 DIQTSQQIQQWADKILEEEXHREAQVSAQLEVHLKQKEOHEYEKIKVLNDQIKKDLA 763
Db 596 KV---TQDNRLQLERLSTLEKE-----CASLELELKATQNKYEQVY-----A 635
QY 764 DIETLENMORHEBAHEKGLISQKAMINAWDSKIRSLRQIRIVELSEANKLAANSILF 823
Db 636 HRET-----EKSRLVSEKXANM---BEVKALQIKLEEKSAQKSDQNS-- 676
QY 824 TQRMKAQEMISELRQOKFYLETQAGLEAQNKLSEQLKISHQDSDKNRLLELTR 883
Db 677 -----QEK---EROISMLSDVYQIQORLOKLEGEYQESKVALHSQ 717
QY 884 LREVSLEHEEQLEKRLQTLQSLQRESOLTALQARAALQSOLQAQTELEFETAE 943
Db 718 IQEQLKXSLQSELGVQ-----RSQTAHLTAREALQVGEVAHLRDAKENVVEELHK 769
QY 944 ABEETQALTAHRDEIQRFDALRNSCTVI---TDLEBQLNQLTEDNAELANNQNFYLSKQ 999
Db 770 LKTAESVDNAQMKELQEQVEAQVFTSLYKTHSNELKEELEKSRHIQEMEERESLVHQ 829
QY 1000 LDEASGANDIEIVQLRSEVDH-----LRREITEREMOLTQKOTMEALKTTCTMLBEQVMDL 1055
Db 830 LQIALARADSEALARSIADESIADELEKETYKLEL-----KELLNKNRTLSQKDISI 883
QY 1056 EALNDELLEKQWE 1070
Db 884 SALRENEQKLE 898

RESULT 14

US-09-005-069-62

Sequence 62, Application US/09005069

Patent No. 5932470

GENERAL INFORMATION:

APPLICANT: FRANK, GLENN R.

APPLICANT: HUNTER, SHIRLEY WU

APPLICANT: WALLENFELS, LYNDIA

TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS

TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,069
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,822
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 900 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-005-069-62

Query Match 9.98; Score 1043.5; DB 2; Length 900;

Best Local Similarity 27.08; Fred. No. 3.7e-46;

Matches 279; Conservative 221; Mismatches 364; Indels 171; Gaps 31;

QY 70 IKVSNFVRXYSDTIAELQELQSAKDFEVRSLVGGCHFAEVQVREKATGDIYAMKVM 129
Db 1 MKSEATNYEIIASSIVNLRMKPDDFNLIKVGAFGEVQLVRHKSTAQVPMKRLS 60
QY 130 KKAALQEQVSPFEERNILSRSTSPWIPOLQVAFQDKNHLVMEEYQPGDLSLLNRY 189
Db 61 KFEIMKRPDGAFFWEERHIMAHAKSEWIVQLHFAFQDKLYVMVMDYMPGGDLVSLMSD 120
QY 190 EDOLDENLIFYLAELTLAVSHVHLGVHRDIKPNILVDRTHGHIKLVDFGSAKMN 249
Db 121 E--IPEKWAYFYTWVVLADTTHSGVHRDVKPNLNLCKYGHKLADFGTCMKMDT 178
QY 250 KMYNAKLPICTPDYMAPEVLTVANGDKGTGYGLDCDWWSVGVIAEMYIGRSPFAEGTSA 309
Db 179 GLVRSNNAVGTDPYISPEVL--QSOGGEVYGRECDWNSVGIPLYEMLFGETPFYADSLV 236
QY 310 RTFNINMNFQFLKFPDDPKVSSDFLDLIQSLLCGQKERL---KPEGLCCHPFFSKIDW- 365
Db 237 GTYSKIMDHNSITPPEVEISQYARSLQGLFDTQELGRNEVEEIKSHFFPINDQWT 296
QY 366 -NNIRNSPPFPVTLKSDDDTGNFDE-----PEKNSWVSSPCQLSPGSGFSGBELPFVG 418
Db 297 FDNLRDSAPPVPELSGDDDTNFDIERDETPEENFPI-----PKTFAGNHLPPVG 348
QY 419 FYSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKE-----LQDSQDKCHKEQ 470
Db 349 FTYNGDYQLLTN-----GGVRSNDMVDTKLNNICVSSKDDVNLQNLLEQKNGSENKLT 403
QY 471 EMTRLHRRVSEVAVLSQKEVELKASQETORSLEQDLATVITCSSLKRSLE---QARME 527
Db 404 NTQLLSNKLDE---LGQRECELR---NQAGDYKELTKFKLSCKELQKAEFENELRK 456
QY 528 VSQ---EDDKALQLLHDIREQSKLOEI-KEQBYQAOVBEEMLMNQLBEDLVSARRSD 583
Db 457 TESLLVETKRLD-----BEQNKRTREMNNOQHNDKINMLEKQINDLQEKL----- 503
QY 584 LYESELRESLAAEFKPKATECOHKLKAKDQKPEVGEYAKLEKINAEQOLKIQLQEQ 643

Db 504 --KGELEHN---QKJQKQAVELR-----VAQSA-TEQLNNELQETWQGLQT 543
Qy 644 KLEKAARERAELEKQNRDSSEGIKKLVABERRHSLKLVKQVLETMRERRLKD 703
Db 544 Q-----RDALQOEVASLQKLSQERSRS---QASDMQIELEAKLQALHIELEHVRNCD 595
Qy 704 DIQTSQOIQOQVADKILEEKEHREQAQVSAQHLEVLHKKOEQYEEKIKVLDNQIKDLA 763
Db 596 KV---TQDNQRLERISTLEKE-----CASLELELKATQNKYQEVK-----A 635
Qy 764 DKETLENMORHEEAEHKGKILSEQKAMINAMDSKIRSLRQIVELSEANLANSSLF 823
Db 636 HRET-----EKSRLVSKKEANM---EYKALQIKNEEKSAQKSDQNS-- 676
Qy 824 TORNNKAQEMISELROCKFYLETQAGKLEAQNKLKEOLEKISHQDHSKXRLLETR 883
Db 677 -----QEK---ERQISMLSDYVQIQOQLKQLEGEYRQESKVALHSQ 717
Qy 884 LREVSLEHEEOKLEKRLQTELQSLQERESQLTALQAAARAALESQLOAKTELETTAE 943
Db 718 IEQQLKKSQQLSELGVQ-----RSQTAHLTAREAQVGEVAHLRDAKRVVEELHK 769
Qy 944 ABEIQAHTARHDEIQRKFDALRNSCTVI-----TDLEEQLNQLTEDNAELNNQNYLSKQ 999
Db 770 LKTARSDVNAQMKELQOEVAEQVFTLYKTHSNELKEELEKSRHIOEMEERESLVHQ 829
1000 LDBASGANDEIVQLRSEVDH---LRREITEREMOLTSQKQWMEALKTTCTMLEBQVMDL 1055
Db 830 LQIALARADSEALARSIADESIADEKXTMKELEL-----KEULNNKRTLSQKDISI 883
Qy 1056 EALNDELLEKEROME 1070
Db 884 SALRERENEQKLE 898

RESULT 15
US-09-171-156A-21
; Sequence 21, Application US/09171156A
; Patent No. 6368846
; GENERAL INFORMATION:
; APPLICANT: Hunter, Shirley Wu
; Weber, Eric R.
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
; APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHERIDAN ROSS P.C.
; STREET: 1560 BROADWAY, SUITE 1200
; CITY: DENVER
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.156A
; FILING DATE: 04-Mar-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Cornell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-171-156A-21
Query Match 9.9%; Score 1043.5; DB 4; Length 900;
Best Local Similarity 27.0%; Pred. No. 3.7e-46;
Matches 279; Conservative 221; Mismatches 364; Indels 171; Gaps 31;
Qy 70 IKHVSFVRKYSDTIALQELQPSAXDFEVSRLVGCGHFAEVQVREKATGDIYAKVMK 129
Db 1 MKSEATYNYETIASVINLRMKPDNFLKIVIGRGAEGEVLVHKSTAQVPAKRLS 60
Qy 130 KKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLYLMEYQPGDGLSLNRY 189
Db 61 KFEMIKRPDSAFWEERHIMAHAKSEWIVQLHFAFDQKLYVMVMDYMPGGDLVLSMSDY 120
Qy 190 EDQLDENLIQFYLAELILAVSHVHLGYVHRDIKPENILVDRTHIKLVDFGSAAKNSN 249
Db 121 E--IPEKAMFYTMVVVLDLTHSMGFVHRDVKPNMLLDKTHGLKLDLADFGCMKMDID 178
Qy 250 KMNNAKPIGTDPYMAPEVLTVMNGDKGTGYGLDCDWSVGVYAIEMYGRSPFAEGTSA 309
Db 179 GLVRSNNAVGTDPYISPEVL--SQSGEGVYGRCDWSVGIIFYEMLFGETPFYADSLV 236
Qy 310 RTFNININFORFLKPPDDPKVSDFLDIQSLICGQKERL---KFEGLCCHPFFSKIDW- 365
Db 237 GTSKINDHNSLTFPPEVEISQVARSILQGFITDRTQRLGRNEVEEIKHHPFFINDQWT 296
Qy 366 -NNIRNSPPPEVPTLKDDDDTSNFDE-----PEKSNWSVSSPCQLSPSGSGEELPFVG 418
Db 297 FDNLRDAPPVVPFELSGDDDTFRNFDIERDETPEENFPI-----PKTFAGNHLFPVG 348
Qy 419 FSYSKALGILGRSESVVSGLDSPAKTSMEKKLLIKSKE-----LQDSQKCHMEQ 470
Db 349 FTYNGDYQLLTN-----GGVRNSDMVDTKLNNICVSSKDDVLNQLNLLQEKGNSENKLT 403
Qy 471 EMTLHRRVSEVAVLSQKVELKASQTSRSLLEQDLATYITECSSLKRSLE---QARME 527
Db 404 NTQLLSNKLDE---IGQRECELR---NQAGDYKEKLTFKLSCKELQRAEFENELRRK 456
Qy 528 VSQ---EDDKALQLLHDIREQSRKLQEI-KEQYQAOVEEMRLMMQLEEDLVSAARRSD 583
Db 457 TESLVLVETKKRLD---EEQNKRTREMNQNNQNDKINMLEKQINDLQEKL----- 503
Qy 584 LYSELSRESRLAAEFKFKATEQCHKLLKAKQDQKPEVGEYKLEKINAEQQLKIQELOE 643
Db 504 --KGELEHN---QKJQKQAVELR-----VAQSA-TEQLNNELQETWQGLQT 543
Qy 644 KLEKAARERAELEKQNRDSSEGIKKLVABERRHSLKLVKQVLETMRERRLKD 703
Db 544 Q-----RDALQOEVASLQKLSQERSRS---QASDMQIELEAKLQALHIELEHVRNCD 595
Qy 704 DIQTSQOIQOQVADKILEEKEHREQAQVSAQHLEVLHKKOEQYEEKIKVLDNQIKDLA 763
Db 596 KV---TQDNQRLERISTLEKE-----CASLELELKATQNKYQEVK-----A 635
Qy 764 DKETLENMORHEEAEHKGKILSEQKAMINAMDSKIRSLRQIVELSEANLANSSLF 823
Db 636 HRET-----EKSRLVSKKEANM---EYKALQIKNEEKSAQKSDQNS-- 676
Qy 824 TORNNKAQEMISELROCKFYLETQAGKLEAQNKLKEOLEKISHQDHSKXRLLETR 883
Db 677 -----QEK---ERQISMLSDYVQIQOQLKQLEGEYRQESKVALHSQ 717
Qy 884 LREVSLEHEEOKLEKRLQTELQSLQERESQLTALQAAARAALESQLOAKTELETTAE 943
Db 718 IEQQLKKSQQLSELGVQ-----RSQTAHLTAREAQVGEVAHLRDAKRVVEELHK 769
Qy 944 ABEIQAHTARHDEIQRKFDALRNSCTVI-----TDLEEQLNQLTEDNAELNNQNYLSKQ 999
Db 770 LKTARSDVNAQMKELQOEVAEQVFTLYKTHSNELKEELEKSRHIOEMEERESLVHQ 829

QY 1000 LDEASGAND EIVQLRSEVDH---LRREITEREMOLTSQKQTMEALKTTCTMLEEQVMDL 1055
Db 830 LQIALARADSEALARSIADESIADLEKEKTKMKELEL-----KELLNKNRTELSQKDISI 883
QY 1056 EALNDELLEKERQWE 1070
Db 884 SALRERENEQKKLE 898

Search completed: July 3, 2004, 10:18:31
Job time : 29 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:07:35 ; Search time 36 Seconds
(without alignments)
5485.589 Million cell updates/sec

Title: US-10-017-216-2
Perfect score: 10490
Sequence: 1 MLKFKYGRNPLDAGAEPI.....QLNGEIRQVQEKSVLRDYC 2053

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7491.5	71.4	1597	2 S68420	citron - mouse
2	1450	13.8	1732	2 T14039	protein kinase (EC
3	1372	13.1	1702	2 T14050	protein kinase (EC
4	1261.5	12.0	1354	2 S74244	serine/threonine-s
5	1253	11.9	1388	2 S76333	serine/threonine-s
6	1250.5	11.9	1354	2 S69211	serine/threonine-s
7	1245	11.9	1388	2 S74245	serine/threonine-s
8	1245	11.9	1548	2 T25808	hypothetical prote
9	1112.5	10.6	1173	2 T25539	hypothetical prote
10	878.5	8.4	624	2 B49364	protein kinase (EC
11	862	8.2	557	2 S71829	serine/threonine-s
12	739.5	7.0	1356	2 T16718	hypothetical prote
13	712	6.8	522	2 G86431	protein kinase T5I
14	706.5	6.7	479	2 S42864	protein kinase (EC
15	706.5	6.7	1256	2 T26101	hypothetical prote
16	700	6.7	441	2 T78393	myotonic dystrophy
17	700	6.7	474	2 T78396	myotonic dystrophy
18	700	6.7	516	2 T78394	myotonic dystrophy
19	698	6.7	756	2 S60966	probable protein k
20	691	6.6	469	2 T41723	serine/threonine-s
21	691	6.6	526	2 S49077	protein kinase PKR
22	687	6.5	569	2 A86170	hypothetical prote
23	687	6.5	596	2 F84589	probable protein k
24	683.5	6.5	1286	2 T14507	hypothetical prote
25	667.5	6.4	443	2 D71405	probable protein k
26	667.5	6.4	475	2 H85156	protein kinase (im
27	665.5	6.3	465	2 I38133	protein kinase (EC
28	661	6.3	500	2 S42867	protein kinase (EC
29	651.5	6.2	483	2 T05188	protein kinase F4I

30	630.5	6.0	665	2 S70706	probable protein k
31	630	6.0	480	2 T47255	serine/threonine k
32	630	6.0	545	2 T01288	protein kinase F27
33	630	6.0	598	2 T47254	serine/threonine k
34	630	6.0	620	2 S22711	probable protein k
35	621.5	5.9	412	2 T78395	myotonic dystrophy
36	599	5.7	908	2 T25035	hypothetical prote
37	594	5.7	1099	2 A56155	tumor suppressor p
38	582.5	5.6	1410	1 A57013	early endosome ant
39	573.5	5.5	2139	2 T18296	myosin heavy chain
40	570	5.4	1909	2 A45592	liver stage anti
41	565	5.4	1738	2 T14867	interaptin - slme
42	559	5.3	2442	2 T08621	centrosome associa
43	541.5	5.2	624	2 T41341	probable serine-th
44	540.5	5.2	2116	2 A26655	myosin heavy chain
45	540	5.1	607	2 S62556	probable serine/th

ALIGNMENTS

RESULT 1
S68420
citron - mouse
C:Species: Mus musculus (house mouse)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C:Accession: S68420
R:Madaule, P.; Furuhashiki, T.; Reid, T.; Ishizaki, T.; Watanabe, G.; Morii, N.; Narumiya
FEBS Lett, 377, 243-248, 1995
A>Title: A novel partner for the GTP-bound forms of rho and rac.
A:Reference number: S68420; MUID:96128238; PMID:8543060
A:Accession: S68420
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1597 <MAD>
A:Cross-references: EMBL:U39904; NID:g1079733; PIDN:AAC52341.1; PID:g1079734
C:Superfamily: protein kinase C zinc-binding repeat homology; pleckstrin repeat homology
C:Keywords: alternative splicing
F:931-979/Domain: protein kinase C zinc-binding repeat homology <X22>

Query Match	71.4%;	Score 7491.5;	DB 2;	Length 1597;	
Best Local Similarity	93.1%;	Pred. No. 2.3e-199;			
Matches 1492;	Conservative 16;	Mismatches 12;	Indels 83;	Gaps 4;	
Qy	468	MEQMTLHRRVSEVEAVLSQKVELKAS	TQSRSLLEQDLATVIT	ECSSLRSLQARME 527	
Db	9	MEQMTLHRRVSEVEAVLSQKVELKAS	TQSRSLLEQDLATVIT	ECSSLRSLQARME 68	
Qy	528	VSQDDKALQLLHDIREQSRKLEIK	EQYQAOVEENRLMMNQL	EEDLVGARRRSDLYES 587	
Db	69	VSQDDKALQLLHDIREQSRKLEIK	EQYQAOVEENRLMMNQL	EEDLVGARRRSDLYES 128	
Qy	588	ELRESRLAAEFKFKATECOHKLK	AKDQKPEVGEYAKLEKIN	AFQOLKIQLQEKLEK 647	
Db	129	ELRESRLAAEFKFKATECOHKLK	AKDQKPEVGEYAKLEKIN	AFQOLKIQLQEKLEK 188	
Qy	648	A-----AKRAERLEKLN	RDSSEGIKKLVEAERH	SLNKKVRL 691	
Db	189	AVKASTEATELLQMIROKAR	ARELEKLNREDSSEGI	KKLVEAERHSLNKKVRL 248	
Qy	692	ETMERRENRLKDDIQTQSK	IQQMADKILEEKHRE	AVSAQHVHLKQKEQHYE	751
Db	249	ETMERRENRLKDDIQTQSK	IQQMADKILEEKHRE	AVSAQHVHLKQKEQHYE	308
Qy	752	KVLNQIKQIADKETLENM	QRHEEAEKGLLSQK	AMINMDSKIRSLQRI	811
Db	309	KVLNQIKQIADKETLENM	QRHEEAEKGLLSQK	AMINMDSKIRSLQRI	368
Qy	812	EANKLAANSLSFTORN	KAQEMISELRQKQF	YLETQAGKLEAQRN	KLEQLEKISHQDH 871
Db	369	EANKLAANSLSFTORN	KAQEMISELRQKQF	YLETQAGKLEAQRN	KLEQLEKISHQDH 428
Qy	872	SDKNRLLETRLRVSL	EHQKLEKRLQIT	ELQSLQERESQLTA	QAARAALESQLR 931

Db 429 SDKRLLETRREVLEHEEOKLEUKQLTELQSLQRESQLTALQARALESQUR 488
QY 932 QAKTELETTAAEETIQAHTARDEIQKFDALRNSCTVITLBERQLNQLTQEDNAELNN 991
Db 489 QAKTELETTAAEETIQAHTARDEIQKFDALRNSCTVITLBERQLNQLTQEDNAELNN 548
QY 992 QNFVLSKOLDEASGANDIEIVOLSEVDHLRREITEREMQLTSQKOTMEALKTCTTMLBEQ 1051
Db 549 QNFVLSKOLDEASGANDIEIVOLSEVDHLRREITEREMQLTSQKOTMEALKTCTTMLBEQ 608
QY 1052 VMDEALNDLLEKKEQWEAWRSVLGDEKSOFCRVREIQRLMDLDEKQGRADQORITES 1111
Db 609 VLDEALNDLLEKKEQWEAWRSVLGDEKSOFCRVREIQRLMDLDEKQGRADQORITES 668
QY 1112 QOVVELAVKHKAEIILALQALKEOKIKABSLSDKLNLEKHAMLENNARSLOQKLETE 1171
Db 669 QOVVELAVKHKAEIILALQALKEOKIKABSLSDKLNLEKHAMLENNARSLOQKLETE 728
QY 1172 RELKORLLESOAKLOQOMDLQKNHIFRLTQGLQALDRADLLKTERSLEYOLENIQVLY 1231
Db 729 RELKORLLESOAKLOQOMDLQKNHIFRLTQGLQALDRADLLKTERSLEYOLENIQVLY 788
QY 1232 SHEKVMKEGTISOOTKLIDFLQAMDOPAKKKGLFSRRKEDPALPTQVPLQYNELKAL 1291
Db 789 SHEKVMKEGTISOOTKLIDFLQAMDOPAKKK-----VPLQYNELKAL 833
QY 1292 EKKARCAELEALOKTRIELRSARSEAAHRKATDPHSTPATARQOIAMSARVSRPH 1351
Db 834 EKKARCAELEALOKTRIELRSARSEAAHRKATDPHSTPATARQOIAMSARVSRPH 893
QY 1352 QPSAMILLAPPSRRKESSTPEFSRRLKERHNNIPHRFNVLGNMRAFKACVCLDTVHF 1411
Db 894 QPSAMILLAPPSRRKESSTPEFSRRLKERHNNIPHRFNVLGNMRAFKACVCLDTVHF 953
QY 1412 GRQASKLEQVCHPKCTCLPATCGLPAEYATHTEAFCDKKNSSPGLQKPESSSLH 1471
Db 954 GRQASKLEQVCHPKCTCLPATCGLPAEYATHTEAFCDKKNSSPGLQKPESSSLH 1013
QY 1472 LEGWMKVPNNKGGQGGWDRKYTVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIH 1531
Db 1014 LEGWMKVPNNKGGQGGWDRKYTVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIH 1073
QY 1532 GAVGASELANTAKA----- 1545
Db 1074 GAVGASELANTAKADVPYILKVESHPHTTWCPRTLYLILAPSPDKQRVTALESWAGG 1133
QY 1546 ---EKAEADAKLLGNSLLKLEGGDRDLDMNCTLPSDQVVLVGTREGLYALNVLNLSLTH 1601
Db 1134 RVSREKAEADAKLLGNSLLKLEGGDRDLDMNCTLPSDQVVLVGTREGLYALNVLNLSLTH 1193
QY 1602 VPGIGAVFIYIKOLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDSPNIFEAV 1661
Db 1194 IPGIGAVFIYIKOLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDSPNIFEAV 1253
QY 1662 KCHLFGACKIENGLCICAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFNYSIL 1721
Db 1254 KCHLFGACKIENGLCICAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFNYSIL 1313
QY 1722 IGTNKFYEIDMKQYTLDEFLDKNDSLAPAVPAASNSFPFVSIVQVNSAGQREBYLLCFH 1781
Db 1314 IGTNKFYEIDMKQYTLDEFLDKNDSLAPAVPAASNSFPFVSIVQVNSAGQREBYLLCFH 1373
QY 1782 EFGVFVDSVGRSRTDILKWSILPLAFAYREPLYFVTHFNLSLEVIEIOARSAGTIPARAY 1841
Db 1374 EFGVFVDSVGRSRTDILKWSILPLAFAYREPLYFVTHFNLSLEVIEIOARSAGTIPARAY 1433
QY 1842 LDIPNRYLGPALISSGAIYLAASYQDKLRVICCKGNLVKESGTEHRRGFSRSPNKRK 1901
Db 1434 LDIPNRYLGPALISSGAIYLAASYQDKLRVICCKGNLVKESGTEHRRGFSRSPNKRK 1493
QY 1902 PPTNEHITKRVASSPAPPEGSHPRESTPHRY--REGRTELRDKSPGRLERKSPG 1959

Db 1494 PPTNEHITKRVASSPAPPEGSHPRESTPHRYVREGRTELRDKSPGRLERKSPG 1553
QY 1960 RMLSTRERSRGLFEDSSRGRLPAGAVRTPLSQVNKRGQSA 2002
Db 1554 RMLSTRERSRGLFEDSSRGRLPAGAVRTPLSQVNKRVQSS 1596
RESULT 2
T14039
protein kinase (EC 2.7.1.37), myotonic dystrophy-associated - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jun-2000
C:Accession: T14039
R:Jeung, T.; Chen, X.Q.; Tan, I.; Manser, E.; Lim, L.
Mol. Cell. Biol. 18, 130-140, 1998
A:Title: Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector
A:Reference number: 217862; MUID:98078670; PMID:9418861
A:Accession: T14039
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1732 <LEU>
A:Cross-references: EMBL:AF021935; NID:g2736150; PID:g2736151; PIDN:AAC02941.1
C:Genetics:
A:Gene: MRCK
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase
F:75-343/Domain: protein kinase homology <KIN>
F:1013-1062/Domain: protein kinase C zinc-binding repeat homology <KZN>
Query Match 13.8%; Score 1450; DB 2; Length 1732;
Best Local Similarity 24.6%; Pred.No.3.4e-33;
Matches 521; Conservative 319; Mismatches 648; Indels 634; Gaps 75;
QY 36 PFMTQQQMSPLSRGILDALFLVFECSSQPALMKIKHVSNFVKYSDDTIAELQELQPSAK 95
Db 18 PAQTNGQC--FSVETLLDILICLYDECNPSLRKKNILLEYLWAKPFTSKVKQMLHRE 75
QY 96 DFEVRSVLGGHFAEVVREKATGDIYAMKVMKKALLAQEQVSFFEEBRLNLSRSTSP 155
Db 76 DFEILKVIKRGAFGEVAVVKLNADKVFAMKILNKWMLKRAETACFBERDVLVNGDSK 135
QY 156 WIPOLQYAFQKQKHLIYMEYQPGDILLNRYEQLDENLIQFYLAELILAVHSVLM 215
Db 136 WITTLHFAFQDDNNLYLWMDYVYGGDILLTLKSFEDRLPEEARFYLAEMVIAIDSVHL 195
QY 216 GYVHRDIKPNILVDRTGHIKLVDFGSAKNGNKNVNAKLPIGTDPYVAPEVLTWANGD 275
Db 196 HYVHRDIKPNILNDMNGHILRADFGSKLMDGTGVQSSVAVGTPTYISPEILQAVE-D 254
QY 276 GKTYGLDCCWNSGVJAYEMIVYGRSPFABGTSAFTFNIMNFORFLKFPDD--PKYSSDF 334
Db 255 GKGRYGECDWNSGLVCMYEMLYGETPFVAESIVETYGKIMHKEKRFQPTQVTDVSENA 314
QY 335 LDLIQSLCCOKERLKEFGL--CCHPFTSKIDWNIRNSPPPEVPTLKSDDDTSNFDEP 391
Db 315 KDLIRLILCSREHRLQNGIEDFKKIPFGSGIDWNIRNCEAPIYIEVSSPTDTSNFDVD 374
QY 392 E---KNSVSSSPQCLSPSGFGSEELPFVGFYSKALGILGRS-----ESV 434
Db 375 DDCLKNSETMPPF--THTAFSGHILPFVGFVTTSVCLSDRSCLRTVAGTSLDLVDNV 431
QY 435 VSGLDSPAKTSSEKXLLIKSKELQSQDKQKQKMEQEMTLHRRVSE---VEA----- 484
Db 432 QRTLDNNLATEYERKI-----KKLEQEKLELTKQESTQTQVALQYSTV 477
QY 485 ---VLSQKEVELKASRTQSLIEQDLATYITESSLSKRSLEQARMEVSGEDDKALQLLHD 541
Db 478 DGPLTASKDLIEKSLKEIEKLKQVA---EVNHLQOOLEAN--SVRELDDAFQIK 532
QY 542 IRECSKLEIKIEQEQVQAQVEENMLNMNOLEDLVSARRSDLYESELESRLAABEFKR 601
Db 533 FEKQIKTLQERE-----ENKELVQASERLKNQSKELKDA----- 568

Db 203 PDNVLLDVNGHRLADFGSCLXNDDGTQSSVAVCTPDYISPEILQAME-DGMSKYGPE 261
Qy 284 CDWWSGVVIAYEIMYGRSPFAGTSARTNNINWFQFLKFPDD-PKYSSPFLDIQSLL 342
Db 262 CDWWSLGVCMYEMLYGETPFYAESLVETYGKIMNHEERFPFPHSHVTDVSEAKDILQRLI 321
Qy 343 CGQKERLKEGGL--CCHPFESKIDANNIRNPPFPVPTLKSDDDTSNPFDEP---KNSW 396
Db 322 CGSERLQNGHEDFKHAFEGELNENIRNLEAPYIPVVSPTSDFVDVDDVLRNIE 381
Qy 397 VSSSPQQLSP---SGFSGELPFVGSYKALGILGRSESVVGLDSPAQTSSMKLLI 453
Db 382 I-----LPPSGHTGFGSLHLPFIGTFT-----TESCFS--DRGSLKSMIQSNTLT 425
Qy 454 KSELODSQDKCHKMEQEMTRLHRRVSEVEAVLSQKEVELKASETCRSILLEQDLATYITE 513
Db 426 KDEDVQDRLNSLOIAYERRR-----449
Qy 514 CSSLKESLEQARMEVS---QEDDKALQLLHDIRE---QSRKLEIKEQYQAVBEMRLM 567
Db 450 -----LEQKLESLRKLOESTQTQSLHGSTRALGNSNRDKEIK--RLNEELERMKS 500
Qy 568 M-----NOLP---EDLVARRSDLYSELSERSLAABEFRRKATECQHKLLKAKDOCKPEV 621
Db 501 MADNSLERQLEDTVTLRQ-----519
Qy 622 GEVAKLEKINAEQOLKIQELQEKLEKAAKERAERALEKLNQREDSSGIRKKLVAEERR 681
Db 520 -----EHEDSTQRLKGLKQYFLAQEKBE-----LHKQLVEASER- 555
Qy 682 HSLNKKVLEWTERENRLKDDIQTKSOIQOMADKILEEKKREBAQVSAHLEVHLK 741
Db 556 -----LKSOTKELKAHQKRALQFSS-----ELNREMAELRSQKQVSRQLR 599
Qy 742 QKEQHYEEKIKVLNDNOIKDLADKETLENMOMRHEBEAEKGIILSEQKAMINAMDSKIR 801
Db 600 DKDEEMEVAMQKIDS--MRQIRKSE-----KSKKELEAR-----632
Qy 802 SLQRIVELSEANKLAANSLFTQRMWKAQEMISELROOKFYLETOAGKLEAQNKLKEE 861
Db 633 -LEDVAEASRKERLREHSESFS-----KQERELETQV---KQGRGPGCATLHQ 681
Qy 862 QLEKISHQDSDKVRILLETRLYESLSHEEQKLELKR-----QLTELQSLQERESQLT 917
Db 682 EISKIRS-----ELEKKV-----LFYEBELVRRERSHVLEKVNKVEHSESHQL 727
Qy 918 ALQAAARAALLESQURAKTELEETTAEEBIEQALTAHRDEIQKFDALRNSCTVITDLBE 977
Db 728 ALQKEVLMKDKLEKSK--RERHSEWEBAIGAM--KQYER-----764
Qy 978 QLNQLEDNALNNQNYLSKQIDELASGAN-----DEIVOLRSE---VDHLREITERENQ 1030
Db 765 ERAMLFDENKLLTAENEKLCFSFYDKLTAQNRQLEDELQDLASKKESVAHWEAQIAETIQW 824
Qy 1031 LTSKOTMEALKTCTWLESOVMDLALNDELLEKQWEAWEVSLGDEKSFQECRVREL 1090
Db 825 VSDEKARGVLOALASQWTE---ELTLRASSLSGSRTLDFLMK--VRSQKLDMSARL-EL 879
Qy 1091 QRMLDTEKQSRARADORITRESRVQVELAVKHEKAEITLAL-QQALKQKLEKSLSKNDL 1150
Db 880 QSALEAE-----IRAKQVHEELRKVKTDTSLAFESKLKESEAK-----917
Qy 1151 EKGHMLENNARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRUTQGLQELADRA 1210
Db 918 -----NRELEEMOSLKRMEEK-----FRADTGL-----942
Qy 1211 DLLKTERSDLEYQLENIQVLYSHEKVMGEGTISQOTKLIDFLQAKMDQPAKKKGLFSRR 1270
Db 943 -----KLPDF-----947
Qy 1271 KEDPALPTQVLOYNELKLALEKEXKARCALEALEALOKTIELRSAREEAHRKATDHPHP 1330
Db 948 -QDPIFE-----YFNAPLAHDTFTTSSASQOETQASKIDL-----SPSV 987

Qy 1331 STPATARQQIAMSIAI VRSPHEQPSAMSL LAPPSSRRKESSTPEFSRRLKERMHNIPHR 1390
Db 988 SVATSTEQQ---EDAARS-QQRESTVPL---PNTQALAMAGPKPKA-----HQ 1028
Qy 1391 FNVGLNNRATKCAVCLD-TVHFGRQASKCLEQCVMHKPCSTCLPATCGLPAYATHFTE 1449
Db 1029 FSIKSPSPQCSHCSTSLMVGLIRQYACEVCAFSCHVSCDSAPQVCPIPE-----1081
Qy 1450 AFCEDKXNSPGLQTKBP-----SSSLHLEGWKMVPRNNKRGQCGQWDRKYIIVLEGSV 1502
Db 1082 -----OSKAPLGVGVORGIGTAYKGVKVPKPT-GVYKGMQRAYAVVCDCKLF 1128
Qy 1503 IYD-NEAREAGQRPVEEFELCPDGVSIHGAVGASELANATAKAEKA-----EADAKLLG-- 1556
Db 1129 LYDLPEGKSTQPGVIASQVLDLRDDEFVAV--SSVLASDVTHATRDIPCFIRVFTASLLGSP 1187
Qy 1557 ---NSLLKL---EGDDR-----LDVN-----CTLPESDOV---1580
Db 1188 SKTSSLLILTENENKRRKVVGLLEGLOAILHKNRLRSOVVHVAQAYDSSLPIKTVLAA 1247
Qy 1581 -----VLVTGEGLYALNVLNLSLTHVPGIGAVFIYIKDLEKLLMIAGERALCLV 1633
Db 1248 AIVDGDRIAVGLEBGLVIELTRDVIIVRAADCKKVQIELAPKEKLILLCGSNH-----1302
Qy 1634 DVKKVQSLOASHLPAPQDIPSPNIFE-----VKGCHLEGAGKIENGLICICAMPS 1684
Db 1303 -----HVHLYPMTSFDGAENSPDIKLPETKGCQLIATGLTRKSSSTCLFVAV 1350
Qy 1685 KWILRYNENLSKYRKETETSEPCS-----CIHF-TNYSILITGNKFEIDM 1732
Db 1351 KRLVLCVEIORTFPFRKFEIIVAPGHVQWAMFKDRLCVGYPGFSLLSIQDGGPLDL 1410
Qy 1733 KQYTLBEFLDKNDHSLAPAVFAASSNSFPVSIIVQVNSAGOREBYLLCFHFEGFVVDYSYGR 1792
Db 1411 -----VNPADPSLA---FLSQQSPDALCAVELKS-----EYLLCFSHMGLYVDPQGR 1455
Qy 1793 RSRTDDLKWSRLPLAPVREPYLFVTHFNLSLEIVETQARSAGTTPARAYLDIENPRYLGP 1852
Db 1456 RSRTQELMWAAPVACSCSSHTVIVSEYGVDFVTRME-----WYQTGLARIRP 1507
Qy 1853 AISSGATYL 1861
Db 1508 LNSDGSNL 1516
RESULT 4
S74244
serine/threonine-specific protein kinase (EC 2.7.1.-) isoform I, Rho-associated - mouse
C:Species: Mus musculus (house mouse)
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jun-2000
C:Accession: S74244
R:Nakagawa, O.; Fujisawa, K.; Ishizaki, T.; Saito, Y.; Nakao, K.; Narumiya, S.
FEBS Lett. 392, 189-193, 1996
A:Title: ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein
A:Reference number: S74244; PMID:9636048; PMID:8772201
A:Accession: S74244
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1354 <NA>
C:Cross-references: EMBL:U58512; NID:g1514695; PIDN:AAC53132.1; PID:g1514696
C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinase; protein kinase homolog
C:Keywords: coiled coil; phosphotransferase; serine/threonine-specific protein kinase
F:74-338/Domain: protein kinase homolog <KIN>
F:1229-1283/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 12.0%; Score 1261.5; DB 2; Length 1354;

Best Local Similarity 26.1%; Pred. No. 4.2e-28; Matches 397; Conservative 289; Mismatches 537; Indels 297; Gaps 51;

Qy 44 SPLREGITDALFVLFPECQPALMKIKHVSNFVKYSYDITAELOEQLOPSAKDFEVSRLV 103

Db 23 SEVNSDCLLDGLDALVYDLDFPALRKKNKTIDNFLSRKYKTINKIRDLRMAEDYEVVKVI 82

```
QY 104 GCCHFAEVQVVRKATGDIYAMKVMKKKALLAQEQVSPFEEERNILSRSTSPWIPQLOYA 163
Db 83 GRGAFGEVQLVRHKSRTKYVYAKMLLSKFEMIKRSDSAFTWEERDIFAFANSPWVQLFYA 142
QY 164 FQDKNHLXWEEYOPGGDILLSLNRYVEDOLDENMLIQVYLAELILAVSHVHMGVYHRDIK 223
Db 143 FQDRDLYVWMEYMPGGDVLNLSNYD--VPEKARFYTAEVVLADALDAHSMGFTHRDVK 200
QY 224 PENILVDRTGHKLVDGSAAXNKNMKNYAKLPIGTDPYMAPEVILTVNMGSGKGTGYLD 283
Db 201 PDNMLLDKSGHLKLADFGTCMKXNKBGMVRCDTAVGTDPDYISPEVLKSQGGD--GYGRE 258
QY 284 CDWWSGVTVAYMIYGRSPFAGTSARTNNIMNQRFKLPFDDPKVSDDFLDLLOSLLC 343
Db 259 CDWWSGVFELYELVGDTPFYADSLVGTYSKIMTHKNSLTFFDDNDISKEAKXNLCFALT 318
QY 344 GOKERLKFEF---LCCHPPFSKID--WNINRSPPPFVPTLKSDDDDTSNFDPEKSNWVS 398
Db 319 DREVLRGRNGVEIEKHLFFKQDQAWETLRDTAVVPDLSDDIDTSNFDLEDKDE 378
QY 399 SS-PCOLSPSGSGBELPVPVGSYKALGILGRSESVSGLSPAKTSMKEKLLIKSKE 457
Db 379 ETFFPI---PKAFVGNQLPVPVGTYSNNRYL-----PSANASENRSNSNVDS 423
QY 458 LODS-QDKCHKMEOEMTRLHRAVSEVAVLSQKEVELKASQTQ---SILFQD-LATYIT 512
Db 424 LQESLQKTIYKLEBQ-----HNEMQLKDEMEQKRTSNLKLKDKIMKELD 468
QY 513 ECGSLKRSLEQARMEVQSDDDKALQLLHDIREQSRKLQEIKEQEYQAOVE-EMRLMMNQL 571
Db 469 EECGNQRNLES----VSQIEKEMLLQHRINERYQKVEQ--ENEKRNINIENSVTLKQDL 523
QY 572 EBDLVASARRSDLYSELESRESLAAEFKPKATECQHKLLKAKDGKQKEVEYAKLEKIN 631
Db 524 E-----DURKASQTSOLANEKUTLOKQLEANDLRLTESHTAVELRKSH 568
QY 632 AFQQLKIQLQELKPKAKERAERLEKLNREDSEGRKIKLVEABEERHSLENKVKRL 691
Db 569 TEMSKSISQL-ESSLNRELQER-NRILENSKSAQDKDYQLQAVLEAERDRG----- 618
QY 692 ETMEERENRLKDDIQTK-----SQOIQMAQDKILEBEKXHRAQVGAQHLVHLKQKBOHY 747
Db 619 -----HDSEMIQDLOARTSLOEEVVKHLKHLNLEVEGERKEAQDVLNISE---KEK---- 666
QY 748 BEKIKVLNDQIKKLADKETTLENMVRHEEAAHE---KGKILSQKAMINAMDSKIRSL 803
Db 667 -----NNLEIDLNYK--LXSIQORLEQFVNEKVTYKARLTDKHQSIIEAKSVAMCEM 716
QY 804 EQRIVELSEANKLAANSLLFTQRNMKAQFEMIS-ELRQCKFYLETQAGKLEAQRNKLSEEQ 862
Db 717 EKKLKEEREAREKAENRVVETEK-----QCSMLDVLKQSQOKLE-----HLTENKERMDE 768
QY 863 LEKISHODSDKNRLELETRLEVSLEHEBQK-----LELKQLTELQLS- 908
Db 769 VKNLALQLEQESNKRLLIQNEUKTOAFADNLKGLKQMKQKEINTLLEAKR-LLEFELAQ 827
QY 909 -----LQERESQTLQARAALESQLRQAKTELEBETTAAEERIEQALTAHRDE 957
Db 828 LTKQYRGNEGQWRELQDLQEAQYFSTLYKTQVKELKEEIEKKNRENLRKIQLQSEKET 887
QY 958 IORKPDALNSCTVITDLEEQNLQNTEDNAELNNQNIFYLSQKLDEASGANDEIVOLRSEV 1017
Db 888 LSTQLD-----LAETKAESQLA--RGILEEQYFELTQESKKAASN----- 927
QY 1018 DHLRREITEREMQLTSQQTWEALTKTCTMLEEOVMQDLEALNDELLEKQWEAARSVLG 1077
Db 928 ---RQEIITDKD-----HTVSRLEETNSVLTQKLEMLRKNKEELNENRRTAESEYKLK 977
QY 1078 DE-----KSQFECRVRELQRLMDTEKQSRARADORITESTRQVVELAVKHEKKAETIALQQA 1132
Db 978 BEEINNLLKAAFE-----KNISTERTLKTQAVNKLAEIMNRKDFKIDRKKA----- 1022
```

```
QY 1133 LKEOKLKAESLSDKLNLEKHALEMNARSLOOKLETERE-LKORLLLEEQAQLQQQMDL 1191
Db 1023 -----NTQDLRKK-----EXENRKLQELAQEREKFNQMVVYHQKELN---DM 1062
QY 1192 QKNHIFRLTQGLQALRADLLKTERSDLEVOLENIQVLYSHKVKMVEGTISQOTKLIDF 1251
Db 1063 Q-----AQIVEEC-----THRNEQLQWLASKESDIEQLEAKLL-DLSDSTSVASF 1106
QY 1252 -----LQAMDQPAKKKKGLFSRRKDEFPALPTQVPLQYNELKLALKEKEKAR 1297
Db 1107 PSADETDGNLPESRIEGLWSVPNRGNIKRYGKKQVYVSVSKLIFYND-----EQDK-- 1159
QY 1298 CAELEBALQKTRIELRSARFEAAHRKATD-----HHPSTPATARQOQIAMSIVRSBEHQ 1352
Db 1160 -----EQSSPSMVLDDIKLFHVRPVTQGDVYR-----AETEEI 1192
QY 1353 PSAMSLPAPSSRRKESSTPEEFERRLKERMHNNIPHRFNVLNMRATKCAVCLDTVHFG 1412
Db 1193 PKIQIILYANEGECRCKDIEVEPVQOGKTNFQNHKGHEFIPITLYHPFANCEACAKPLWHV 1252
QY 1413 ROASKLCLEQVMCHPKC-----STCLPATCGLPABYATHFTFAFCORDKXNSGLQTK 1465
Db 1253 FKPPPALECR-RCHVKCHRDHLKKEDLISPKVSYDVTS-----ARDMLLLACSDEQ 1305
QY 1466 PSSSLHLEGMKVPNNKRG 1485
Db 1306 KKWVTHLV--KKIPKNPPSG 1323

RESULT 5
S70633
serine/threonine-specific protein kinase (EC 2.7.1.-), Rho-associated - bovine
N/Alternate names: Rho-associated protein kinase
C/Species: Bos primigenius taurus (Cattle)
C/Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 24-Sep-1999
C/Accession: S70633; S77694
R/Matsui, T.; Amano, M.; Yamamoto, T.; Chihara, K.; Nakafuku, M.; Ito, M.; Nakano, T.; O
EMBO J. 15, 2208-2216, 1996
A/Title: Rho-associated kinase, a novel serine/threonine kinase, as a putative target for
A/Reference number: S70633; MUID:96208507; PMID:8641286
A/Accession: S70633
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1388 <MAT>
A/Cross-references: EMBL:U36909; NID:g1326077; PIDN:AAC48567.1; PID:g1326078
A/Accession: S77694
A/Molecule type: protein
A/Residues: 1-18;30-34;36-44;58-64;133-140;248-252;291-295;297-305;327-347;350-360;366-37
-1070 <MAT2>
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase C zinc-
C/Keywords: Atg; phosphotransferase; serine/threonine-specific protein kinase
F/90-354/Domain: protein kinase homology <KIN>
F/98-106/Region: protein kinase ATP-binding motif
F/1261-1315/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 11.9%; Score 1253; DB 2; Length 1388;
Best Local Similarity 26.1%; Pred. No. 7.3e-28;
Matches 394; Conservative 291; Mismatches 507; Indels 320; Gaps 50;

QY 15 GAPEPTA---SRASNLNLFQGGPPFWTQQQMSPLSRGILLDALFVLFECSSQPALMKIK 71
Db 12 GAPEAVSGGAGASR-----QRKLEALIRDPSPINVESILLDGLNPLVLDLDFPALRKXK 66
QY 72 HVSNFVKYKSDTIAELQELQPSAKDPEVRSVLCGCHFAEVQVVRKATGDIYAMKVMKK 131
Db 67 NIDNFLNRYEKIVKIRGLQMKAEEDVVKVIGRGAFGEVQLVRHKASQKVYANKLLSKF 126
QY 132 ALAQSQVGFEEERNILSRSTSPWIPQLOYAQDQKNHLYLMEEYOPGGDLSLNRYED 191
Db 127 EMIKRSDSAFFWEERDIFAFANSPWVQLFCAFDQDKKLYLMVMEYMPGGDVLNLSNYD- 185
QY 192 QLDENLIQVLAELILAVSHVHMGVYHRDIPENILVDRTGHIKLVDGSAKKNMKNM 251
```

186 -VPEKWAFTAEVVALDAIHSMGLIHRDVKPDNNLLDKHGHKLADFGTCKMKMDTGM 244
Db
252 VNAKLPIGTPDYVAPVLTVMGDKGTGTLGDCDWSVGVYIAYEMIGSPAEQTSART 311
Qy
245 VHCDDTAGTDPDVISPEVLKSGGD--GYGREGCDWMSVGVFLFEMLVGDTPPYADSLVGT 302
Db
312 FNIMNFORFLKPPDPKPYSSDFDLIOQLCCQKRLKFEQ---LCCHPFPFSKIDWN--- 366
Qy
303 YKIMDHKNSLCEPDEABEIKKAKNLICAFIDREVLGRNGVEBELKQHPFFKNDQWMD 362
Db
367 NTRNSPPVPTLKSDDTSNDEBPKNW-VSSPCQLSPSGFGEELPFVGFYSKAL 425
Qy
363 NIRETAAPVVPVPELSSDIDSNFDDIEDDKGVETETPI---PRAFGVGNQLPFGFTYYREN 419
Db
426 GILGRSESVVGLDSPA--KTSMEKKLIIKSKELQDSODKCKMEQEMTRLHRRVSEVE 483
Qy
420 LLUS-----DSPCKENSIQSR---KNEESQEIQKLYTLEHLS----- 457
Db
484 AVLSQKEVELKASQTSRSLLEQDLATYITECSLSKSLSEQARMEVSYQEDDKALQLLHDIR 543
Qy
458 -----TEIQAKEE---LEQ-----CKSVNTRLEKVAKELEBEITLRKNVESTLR 499
Db
544 --FQSRKLOEIKQEQYQAEVQVNRMMNOLBEDLYSARRSDLYSSELESRLAAEEFKR 601
Qy
500 QLEREKALLQHKNAEYORKADHEADKKRNLENDVNSLKQDLBKRNQNSQISTE---- 555
Db
602 KATPCQHKLLKAKDQKPEVGEVAKLEKINAEQKIQ-----ELQEK--LEKAAKER 652
Qy
556 KYNOLQRLQDETNALLRTESDTAARLRTQAESSKQIQLESNNRDLQKNCLETAKLK 615
Db
653 ARELEKLNQNRDSSGIGIKKLVABERRHSLNKKVRLTWERENRDKDQIYKSOQI 712
Qy
616 LEKEFINQSVLESE---RRDRTHGSEIINDLQGRISGLEEDVYQNGKILLAKVELEKQL 672
Db
713 QOMADKILEEKEHREAOVSAQH---LEVHLKQKE-QRYEEKIKVLD-NQIKKDL--AD 764
Qy
673 Q---ERFTLEKKNWMDMTYQVKVIOQSLEQETEHKATKARLADCKIVSEBAK 729
Db
765 KETLENMQHREBEAHEKGI---LSEQKAMINAMDSKIRSLQRIVELSE-----ANKLA 817
Qy
730 SEAMKEMKLSBERTLKQKVENLLAEKRCRSLDCDLKQSQKINELLKQDVLNEDV 789
Db
818 ANSSL-----FTORMKQAQEWISLRLQKQKPYLETQAGLEAQNRLKEOLEKIS 867
Qy
790 RNUTLKIEQTKRCITQNDLRKQGTQOVNFKWS-----EKQLKQ-- 829
Db
868 HQDHSKRNRLLETRILREVSLEHEQKLEKRLQTLQLSQRESQLTALQAAALE 927
Qy
830 -----ENNHLKMSL-----EKQNAELKRCQDQADQMKQLQDLQEAQYESTLYK 877
Db
928 SQRQAQTELEBTT---AEABEIQALTAHREIQRKFPALRNSCTVITDLBQLNOLTE 984
Qy
878 TOVRELKEECEETKCKELQKQKQELQDBRDSLAQLE-----ITLTAKDSEQLAR 929
Db
985 DNEALNNQNFYLSKQDASGANDVQLRSEVDHLRRIITEREMQLTQSKQTMELAKTT 1044
Qy
930 STAEQYSDLEKIMKE-----LEIKEMWARKHQLTEKDAIASLEETNRLTSD 981
Db
1045 CTMLEQVMDLEALNDELLEKQEWANRSLVQDE-----KSQFECVRLEQLRMLDTE 1097
Qy
982 VANLAN---EKEELNKLKEAQEQ---LSRLKDEEISAAAIKAQPE-----KQLLTE 1027
Db
1098 KQSRARADORITESTRQVVELAVKHEKAILALQOALKQKLEKESLSDKINDLEKHAML 1157
Qy
1028 RTUKTQANKL-----AEINNRKEPVK-----RGNDTVERK-- 1059
Db
1158 EMNARSQQKLETERELKORLLEBQAKLQOQMDLQKRNHFRLTQ---GLQBALDRADLLK 1214
Qy
1060 EKENRKLHMBELKSERE---KLTQOMIKYQKELNEMQAIAEESQIRIELQMTLDSKD--- 1113
Db
1215 TERSDLEYQLENTQVLY---SHEKVKVQEGTIS---QOQTKLIDF 1251
Qy
1114 ---SDIEQLRSQALHIGLSDSSIGSGPGDTEADGFPESRLGLEWLSLPVRNNTKKFGH 1170
Db

1252 LOAKMDQPAKKKKGLFSRRKEDPALPTQV-----PLOTNELKLALEKEKARCAEL 1301
Qy
1171 VKXVIVSSKKILFYDSEQKQSNFVWLIDKLFHVRPVTQTDVVRADAKEIPRIFQI 1230
Db
1302 EALQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAP 1361
Qy
1231 -----LYAN 1234
Db
1362 PSSRRKESSTP---BEFSRRLKERMHNIPHRFNVGLNMRATKCAVCLDVTVHFGQASK 1417
Qy
1235 EGSKKQEQEPFVPEVGEKSNYICHKGHEFTPLYHF-----PINCEACMKPLMHMFKPPP 1289
Db
1418 CLEQVNCCHKPC 1429
Qy
1290 ALECR-RCHIKC 1300
Db
RESULT 6
S69211
serine/threonine-specific protein kinase (EC 2.7.1.-), Rho-associated - human
C:Species: Homo sapiens (man)
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 24-Sep-1999
C:Accession: S69211; S71910
R:Ishizaki, T.; Maekawa, M.; Fujisawa, K.; Okawa, K.; Iwamatsu, A.; Fujita, A.; Watanabe,
EMBO J. 15, 1885-1893, 1996
A:Title: The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr prote
A:Reference number: S69211; MUID:96203110; PMID:8617235
A:Accession: S69211
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1354 <ISH>
A:Cross-references: EMBL:U43195; NID:g1276900; PID:AB02814.1; PID:g1276901
A:Accession: S71910
A:Molecule type: protein
A:Residues: 187-195;281-288;465-473;573-587;818-828;885-893;934-945 <ISH2>
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase C zinc-
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:74-338/Domain: protein kinase homology <KIN>
F:182-90/Region: protein kinase ATP-binding motif
F:1229-1283/Domain: protein kinase C zinc-binding repeat homology <KZN>
Query Match 11.9%; Score 1250.5; DB 2; Length 1354;
Best Local Similarity 26.0%; Pred. No. 8.4e-28;
Matches 393; Conservative 298; Mismatches 543; Indels 275; Gaps 51;
Qy 44 SPLSRGILLDALFVLFEECSQPALMKIKHVSFNRKYSQDTIAELQELQPSAKQPEVRSLV 103
Db 23 SEVNSDCLDGLDALVYDLDLFPALRKNKINDNLSRYKOTINKIRDURKAEDEVVKVI 82
Qy 104 GCGHFAEVQVREKATGDIYAMKMKKALLAQEQVSFFFEERNILSRSTSPWIPQYQA 163
Db 83 GRGAFGEVQLVRHSTKRVYANKLLSKFEMKSDSAFFWEERDIMAFAFANSPWVQLFYA 142
Qy 164 FQDKNHYLYMEYOPGGDLSLNRYSDDQDENLIQYLAELILAVSHVLMGVHEDIK 223
Db 143 FQDDRYLYMVMYMPGGDLVNLANSYD--VPEKWARFYTAEVVLADAIHSMGFIRHDVK 200
Qy 224 PENILVDRTHGHIKLVDFGSAAKNNKNNVNAKLPIGTPDYMAPEVLTMNGDKGTGYLD 283
Db 201 PDNMLLKSHGLKADFGTCMKNNKEGVRCDAVAGTDPYISPEVLKSGGD--GYIGRE 258
Qy 284 CDWMSVGVYAVEMTYGRSPFAEGTSARTFNINMNFQRLKFPDDPKVSSFDLIQSLLC 343
Db 259 CDWMSVGVYVEMLVGDTPPFYADSLVGYSKIMHKNLSLTFPDDNDISKEAKNLICAFLT 318
Qy 344 QKRELKFEQ--LCCHPFFSKID--WNINRNSPPFPVPTLKSDDDTSNDEPEKNSWVS 398
Db 319 DREVLGRNGVEETKRLHLPFKNDQAWETLDTVAVPVVPLSSDIDTSNDDLEEDKGE 378
Qy 399 SS-FCQSPSGFSGEELPFVGFYSKALGILGRSESVVGLDSPAKTSMEKLLIKSKE 457
Db 379 ETFFI---PKAFVGNQLPFGVFTY-----YSNRRYLLSSANPNDNRTSSN-----ADKS 423

QY 458 LQDS-QDKCHKMEQETRLHRRVSEVEAVLSQKEVELKASQETOR---SILEOD-LATYIT 512
Db 424 LQESLQXTYKLEBQL-----HNEMQLKQMEQKCHTSNIKDKWKMLD 468
QY 513 ECSSLKRSLEQARMEVQSDBDKALQLLHDIRBOSRKLQEIKEQEQYQAVE-EVRLMMNL 571
Db 469 EECNQRNLEST---VSQIEKEMQLQHRINEYQKAEQ--ENEKRNVENEVSTLKQQL 523
QY 572 EEDLVARSARDLYESLRESLAAEFKPKKATECOHLLKAKKQDQKPEGEYAKLEKIN 631
Db 524 -EDLKKVNSQL-----ANE---KUSLOQLEANEANDLLRSTDAVLRKSH 568
QY 632 AEQQLKIQELQEKLEAKKERAERLEKLQNRDSESGIRKLVAEABERHSLNKKVRL 691
Db 569 TEMSKSISQL-ESLNRELQER-NRILENSKSTQDKDYQLQALAEARDRG-----618
QY 692 ETWERENLKODIQTK-----SQIQCMADKILEEKEHREAVQVSAHQHLEVHLKQEQHY 747
Db 619 -----HDSMIGLOARITLSQBEVHLKXNLKVEGERKEAQDMLNHE---KEK---666
QY 748 EEKIKVLDNQIKDLADKETLNNMORHEEEAHE-----KGKILSEOKAMINAMDSKIRSL 803
Db 667 -----NNLEIDLNYK--LKSLOQLRLEQVNEHVKTKARLTDKHQSIEBAKSVAMCEM 716
QY 804 EQRIVELSEANKLAANSLSFTORNKKAQEMIS-ELRQOKFYLETQAGKLEAQNRKLEQ 862
Db 717 EKKLEEREAREKAENRVV---QIEKQCSMLDVLKQSQQLLEHLTGKE---RMEDE 768
QY 863 LEKISHQSDKRNLELETRLEVSLEHEEOK-----LELKROLTELOLS- 908
Db 769 VKNLTLOEQENKRLLLQNEIKTQAFADNKLKLEKQKQINILLKAKR-LLSEFEAQ 827
QY 909 -----LORESQTLQAAPAALESQRQAKTELETTAAEIEIQALTAHRDE 957
Db 828 LTKQYRGNEGQRELQDQEAQYFSTLYKTQVKELKEIEBKRENLENKIKIQELQNEKET 887
QY 958 IQRKEDALRNSCTVITDLEEQNLQNTEDNAELNNQNFLYSKQDLQDASGANDEIVOLRSEV 1017
Db 888 LATQLD-----LAETKAESEQA--RGLLEEQYFELTQSSKKAASRN-----927
QY 1018 DHLRREITERMQLTSQKQMBALKTTCMLFEEQVMDLEALNDELLEKEROEAWRSVLG 1077
Db 928 ---RQEIITDKD-----HTVSRLEANSMLYKDIEILLRRENEEITKXKKAEEYKLEK 977
QY 1078 DE-----KSQFCRVRELQMDLTKQSPARADQRTSQRQVVELAVKHAELALQQA 1132
Db 978 EEEISNLKAAFS-----KNINTERLTQAVNKLAEIMNRKDFKIDRKKANTQDLRKK 1030
QY 1133 LKEQK---LKAESLSDKLNDEKKHAMLEMNARSLOQKLETERELKORILLEQA---KLO 1186
Db 1031 EKENRKLQLELNQEREKFNQVVKH-----QKELQ---NDMQAQLVECAHNRNELQ 1077
QY 1187 QQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENTQVLYSHEKVKVQEGTISQQT 1246
Db 1078 MOLASKESDIEQLRAKLLDSDSTSVASPPSAD-----ETDGNLPESR 1120
QY 1247 KLIDFLQAKMDOPAKKKGLFSRRKEDPALPQVPLQVYNELKLALEKARCAELEELAQ 1306
Db 1121 -----IEGWSLVFNQGNITKRYGKKQYVWVSKLIFYNEDQCKEQSNPWSVLDIDKL- 1173
QY 1307 KTRIELRSAREBAARHKATDHPHPSTPATARQOIAMSAIVRSPEHQPSAMSLAPPSSRR 1366
Db 1174 ---FHVPRVPTQGVYRAET-----EETPKIIFQILYANEGEC 1206
QY 1367 KXSSTPEPSRRLKERMHNIHRRNVGNLNMATKCAVCLDTHVHFGRAQSKLCQGVNCH 1426
Db 1207 RDXVEMEPVQQAETNFQNHKGHEFIPTLYHFPANCDAKAPLMHVFVPPPALECR-RCH 1265
QY 1427 PKCS-----TCLPATCGLPAEYATHFTFAFCRDKMNSPGIQTKEPPSSSLHLEGW 1476
Db 1266 VKCHRDHLDKKEDILIC-----PKVSYDVYTSA--RDMILLACSDQCKKQVTHLV--K 1314

QY 1477 KVERNNKRG 1485
Db 1315 KIPKNPPSG 1323
RESULT 7
S74245
serine/threonine-specific protein kinase (EC 2.7.1.1-) isoform II, Rho-associated - mouse
C:Species: Mus musculus (house mouse)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 24-Sep-1999
C:Accession: S74245
R:Nakagawa, O.; Fujisawa, K.; Ishizaki, T.; Saito, Y.; Nakao, K.; Narumiya, S.
FEBS Lett. 392, 189-193, 1996
A>Title: ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein s
A:Reference number: S74244; MUID:96368048; PMID:8772201
A:Accession: S74245
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1388 <NAK>
A:Cross-references: EMBL:U58513; NID:g1514697; PIDN:AAC53133.1; PID:g1514698
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase C zinc-
C:Keywords: phosphotransferase; serine/threonine-specific protein kinase
F:90-354/Domain: protein kinase homology <KIN>
F:1261-1315/Domain: protein kinase C zinc-binding repeat homology <KZN>
Query Match 11.9%; Score 1245; DB 2; Length 1388;
Best Local Similarity 25.8%; Pred. No. 1.2e-27;
Matches 381; Conservative 296; Mismatches 517; Indels 284; Gaps 45;
QY 32 QGKPPFTQQOMSPLSREGILDALFVLFECSPALMKIKHVSNFVKYSDTTAELQELQ 91
Db 27 QRLKALIRPRSPINVSLELDGLNSLVLDLDFALFKKNKIDNLFNRYKIVKIRGLQ 86
QY 92 PSAXDFEVRSLVGGCHFAEVQVVRKATGIYAKVMKKKALLAQEQVSPFEERNILSR 151
Db 87 MKAEDYDVXVVGGA-GEVQLVRHKASQVYAMKLLSKEMIKRSDSAFFWEERDIMP 146
QY 152 STSWIQLQVAFQDKXNHLIMEEYQPGDLISLLNRYEDQLENLQFYLAELILAVHS 211
Db 147 ANSWVWVQLFCAFODDRYLYVMVMEYMPGDLVNLMSYD--VPEKMAKFTYAEVVALDA 204
QY 212 VHLNGYVHRDIPKPNILVDRTHIKLVDFGSAKNSKVMNAKLPITGTPDYMAPEVLTV 271
Db 205 IHSGLIHRDVKPNMLDKHGLKLAFTGTCWKNDTGMVHCDTAVGTTPDYISPEVLKS 264
QY 272 MNGDGKTYGLDCDWVSGVIAYEMIVGRSPFAEGTSARTFNIMNFQRLKPPDPKVS 331
Db 265 QGGD--GYGREGCDWWSGVFLFEMLVGDTPFVADSLVGTYSKIMDKNSLCPDETEIS 322
QY 332 SDFLDLITQSLCCQKELKTEG---LCCHPFSKIDWN--NIRNSPPFPVPTLKSDDDTS 386
Db 323 KHAKNLICAFLTDFREVLGRNGVEETKQHPFFKXNDQNMNDNIRETAAVVPVPELSSDIDSS 382
QY 387 NFDPEKNSW-VSSSPQQLSPGSGFSGEELPFVGFYSKALGILGRSESVVSGLDSP--AK 443
Db 383 NFDIEDDKGDETFFI---PKAFVGNQLPFIQFTYFRENLLS-----DSPCRE 430
QY 444 TSSNEKLLTKSELQDSQDKCHKMGEMLHRRVSEVEAVLSQKEVELKASQETORSL 503
Db 431 NDAIQTR---KSEESQBIQKKYALEEHL-----SEVQA---KEELEOK-----469
QY 504 EQDLATVITCSLSKRSLEQARMEVSGEDDKALQLLHDIR--EBSRKLQEIKEQEQYQAV 561
Db 470 -----CKINTRFLEKTALEETILRSVSTLRQLEREKALLQHKNAEYORKA 519
QY 562 EEMRLMMNLQLEBDLVARSRRSDIYSELRESLAAAEFFKRAKATECOHLLKAKKQKQKPEV 621
Db 520 DHEADKRNLENDVNSLKQLEDLKKRNQSSQISTE---KVNQLQKQLDEANALLRTES 575
QY 622 GEYAKLEKINAEQOLKIQ-----ELQEK--LEKAAKERAERLEKIQNRDESSSEGIRK 672
Db 576 DTAARLRKTOAESKQIQQLLESNNRDLQDKCULETAKUKLEKEFINLSALES-----629


```
Db 671 IKLKL-----QRMDERASHTAQSEQENKQLEAHYERAKM-----L 707
QY 976 EEQLNLTBDNAELNNQNFYLSKQLD--BAGANDEIVQLRSEVDHLRRE-ITEREMOLT 1032
Db 708 QDNVQMVAVENRGLRDEIEKLQSQMAALPRGGLNEQ--QLHEIFNVVSEKATREMEMNL 765
QY 1033 SQKQT--MEALKYTCMLSEQWMDLEALNDELLEKRWQWAVRSVLGDEKSFQRCVREL 1090
Db 766 TRKITGEVSLKNNSLPTTSNYI-----QNTPSGWS-----RRMNV 803
QY 1091 QRM--LDTEKQSRARADORITESQVVELAVXEHKAEILALQALKEOKLKAESLSDKLN 1148
Db 804 ARKGLDLQRLQAEIDAK-----LKLKAEILKNSGEQVLTSAARDL 844
QY 1149 DEKKHAMLENNARSLOQKLETERELKORLLEQAKLQOQMDLQ--KNHIFRLTQGLQEA 1206
Db 845 DTEKRWMA-----SLMREVAMLKQKNIENSDDSAFSTWG----- 879
QY 1207 LDRADLLKTERSDLEYOLENIQVLYSHEKVXVEGTIS-----QOTKLIDFLQAKM 1256
Db 880 --RGDLMISWND--YEMSN--SSLRQEMISQSTPSYENALLHDHQVPRKVDLRYK- 933
QY 1257 DOPAKKKGLFRRKXEDPALPTQVPLQYNELKLALEKAKARCAELEALQKTRIELRSAR 1316
Db 934 QXPMKTASGIFS-----TKCGHCTSILGLDRQGLFCQSQVACHVSCAERSVQ 1003
QY 1317 EEAHRKATDHPSTPATARQOIAMSAIVRSPHQPSAMSLAPPSSRRKESSTPEBFS 1376
Db 946 -----PVSTISAMERG--HNFERMKIKTP----- 966
QY 1377 RRLKERMHHNI PHRFNVGLNMRATKCAVCLD--TVHFGROASKCLCQVMCHPKSTCLPA 1435
Db 967 -----TKCGHCTSILGLDRQGLFCQSQVACHVSCAERSVQ 1003
QY 1436 TCGLPAYATHTEAPCRDKMNSPLQTKPESSSLHEGWMKVPNNRKGQOGWDRKXIV 1495
Db 1004 SCVPTEE-----ERRPLGMIRH--- 1020
QY 1496 LSGSKVLIYDNAREAGQRPVEFELC-LPDGDVSIHGAVG-----ASE 1538
Db 1021 AEQNMQDVKNRILRLVMDRDPDFTVCGVSEADV-IHAQKGDIPIKIFRVTITQILNSSF 1079
QY 1539 LANTAK-----AEKAEADAK-----LLGNS-----LLKLEGD-----DRL 1568
Db 1080 YSSSKFYTLFAETEERKKVWALSELKTLRRSKLADRKAFLVKEVDFVTLPLSRV 1139
QY 1569 DMNCTLPFSDQVLYGTEGLVALNVLKNSLTHVPGIGAVFIYIILKLEKLLMAGGER 1628
Db 1140 AQCCAIIDRSKIVGFSHGGLYCIETISROLLIPVG-----EK-----ENQR 1182
QY 1629 ALCLVDVKVKQSLAQSHLPAQP-----DISPNIFEAVKGCGLFGAGKIE 1674
Db 1183 CVETVEYDEAEQLLMMIVGPAKDRHVRIVPSAALDGRDLKWKVNDTKGCHLLAVGTNVP 1242
QY 1675 G---LCICAMPKSKVILRYNLSKCYCKRIEITSEPCSCITHFNYSILIGTNKFYEI- 1730
Db 1243 GGRAGFFAVAFKSVTIFQIDRSEKRRKKWKDLAMPGTPQSTAIENGRLYVGFHSFRS 1302
QY 1731 -----DMKQYTL--EFLDKNDHSLAPAVFAAASSNPVSIVOVNSAGQREE 1775
Db 1303 SLVGVDSPPVSGDASGAVLQHSILVNMDTSLO---FLNQOTSTEAKLI-VNVPSPDE 1358
QY 1776 YLLCFHFQVVDVSGRRSRDLDLKWSLPLAPAYREPLYFTHFNLSLEVII 1828
Db 1359 YLLVFNMIGLYVNMGRSRLPEVFPPTQAKYFAYHEPILCVFSENEVDIFNV 1411
```

RESULT 9

T25539

hypotheoretical protein C10H11.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000

C:Accession: T25539

R,Dante, M.; Wamsley, P.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C10H11.
A:Reference number: Z20047
A:Accession: T25539
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1173 <DAN>
A:Cross-references: EMBL:U08311; PIDN:AAB42348.1; GSPDB:GN00019; CESP:C10H11.9
A:Experimental source: strain Bristol N2; clone C10H11
C:Superfamily: hypotheoretical protein C10H11.9; protein kinase homology

```
Query Match 10.6%; Score 1112.5; DB 2; Length 1173;
Best Local Similarity 29.9%; Pred. No. 4.6e-24;
Matches 291; Conservative 210; Mismatches 346; Indels 127; Gaps 33;

QY 44 SPLSRGILDALFVLFECCOPALMKIKHVSFVRKYSDTIAELOELQPSAKDFEVSILV 103
Db 15 SPINIESLDTITALTVDCKIPVLMRMKSVDFNISRYERVVESLALRMKAADFQRLKVI 74
QY 104 GCGHFAEVQVREKATGDIYAMKVMKKALLAQBOVFFEEERNILSRSTSPWIPQIQA 163
Db 75 GRGAFGEVHLVRHTRTNTVYAMKVLNKDDMIKRADSAPFWEERDIMAHANSEWILRLOYA 134
QY 164 FQDKNHYLVEEYOPGGDLLSLNRYEDQDENLIQYLAELILAVHSVHLMGYVHRDIK 223
Db 135 FQDPRLHYMVEYMPGGDLVNLATSYE--VSEKWTREYTAIVEALAAALHSMGYIHRDVK 192
QY 224 PENILVDTGHIKLVDFGSAKKNKMNKNAKLPIGTPDYMAPEVLTVMNGDGKGTGLD 283
Db 193 PDNMLIISIGHIKUADFGTCVMNANGVVRCSCTAVGTFDVISPEVLRNQQDAB--FGKE 250
QY 284 CDWMSVGVIAYEMITGRSPFAECTSARTFNINMNFQFLKFPDDPKVSSDFLDLIQSLLC 343
Db 251 VDWMSGVGF--YEMLVGETPFVAEALVSTYTNIMNKHSLAKFPDPLISTQAKDIKKFLS 310
QY 344 GQKERL---KFEGLCCHPFFSKIDWN--NTRNSPPPPVPTLKSDDDTSNTDEPEKNSWVS 398
Db 311 AAPRLGRNSVDDIRNKHFFVNDWTEPATLEASPPVIPSLSKSDDDTTFEEIETRDRDN 370
QY 399 SSPQLSPSFGSBEELPFVGFYSKALGILGRSESVVSGLDSPAKTSSME--KKLLKSK 456
Db 371 AGDFQL-PKTFNGNLDPFIQFTYSNEYSVP---KNLLKHGAGSKQNGIEQHKPQTVEQ 426
QY 457 ELQDS-----QDKCHRMQEMTRLHRRVSEVAVLQSKVELKASTORSLLLEQDLATY 510
Db 427 PLTNHASGVPEEKYEAVKMLDPSKNEFELLKDSIARNEIRAKMIENKNSLS----- 480
QY 511 ITECSSLKRSLEQARMEV---SQEDDKALOLLHDIREQSKLOEIKEQEYQAOVEEMELM 567
Db 481 -TKISDLERELKONKOLRHGADSDAKVNELAVELR-----MSKEYNSESELSKFRDK 534
QY 568 MNQLEEDLVARSARDIYSELRESRL-----AAEFKPR-----KATECOHKL 610
Db 535 CEQLKEDL---RKXSGELAQEKNETQVFPQKXDADEAFAEIKRDYELLQTRNEKSVQL 591
QY 611 LKAKQCKPEVGEYAK--LEKINAEQOLKIQELQEKLEKAAKABARELEKIQNRESSE 668
Db 592 KKALDERK-ENGAYQOSVAKATDAEWKQVQFYEQLEHANDER-KREEOKRTAAEPDQS 649
QY 669 GIRKKLVEABERRHSLNKVKRL-ETMERRENKDKDIQTKSQ---QIQOMAD-----K 718
Db 650 RVARKLAGIEANYELLQNDYKSMKEARKDLERDQDVITEKRRLRIRVEQLMDSRNTDER 709
QY 719 ILEL-EKKHREAQVSAOHLVHLKQK---BOHYEEKIKVLNDQIKKDL--ADKE----- 766
Db 710 VLSLQDELVESQFEAKYKEDGLRGIDGFKHELENEKMKQTLEENLLVADKRGKLMK 769
QY 767 TLENMQRHEEAHEKGIKLLSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLTQR 826
```



```

Db      770 EVQELMQRKWEITNK-----DQTLKHELTQDLEIKQSKIESSEQ-----ES 812
Qy      827 NMKAQEMISELQQKXLEYLTQAGKLEAQNKRLEEGOLEKISHQDSD----KNRLLELET 882
Db      813 NDK---QTIADLRKK---LDLEAKHKAVINKLEEMAKRQPLKKGKGVTSALTKKER 866
Qy      883 RLREVSHEBEQKL-----ELKRLQTELOLSQERESQITLQAARAALESQLOQA 933
Db      867 EI--MALEQDRTMSKRIALFAYENDKQAEHFNIAQD-----NQTQDALRELKEC 917
Qy      934 KTBLEETAEABEE 947
Db      918 KEELANRNVNTRYE 931

RESULT 10
B49364
Protein kinase (EC 2.7.1.37), myotonic dystrophy-associated - human
N;Alternate names: myotonin protein kinase
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text change 24-Sep-1999
C;Accession: B49364; S48681; A42101
R;Shaw, D.J.; McCurrach, M.; Rundle, S.A.; Harley, H.G.; Crow, S.R.; Sohn, R.; Thirion,
Genomics 18, 673-679, 1993
A;Title: Genomic organization and transcriptional units at the myotonic dystrophy locus.
A;Reference number: A49364; MUID:94140369; PMID:7905855
A;Accession: B49364
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-624 <SHA>
A;Cross-references: GB:119268; NID:G307177; PIDN:AAA36206.1; PID:G307177
R;Sasagawa, N.; Sorimachi, H.; Maruyama, K.; Arahata, K.; Ishiura, S.; Suzuki, K.
FEBS Lett. 351, 22-26, 1994
A;Title: Expression of a novel human myotonin protein kinase (MTPK) cDNA clone which enc
A;Reference number: S48681; MUID:94357271; PMID:8076686
A;Accession: S48681
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-417, 'L', 419-468, 'P', 470-544, 'MAPRPWLWASRWGQAPCTAATCCSLPGS', 'LGIAVRRRPCCS
A;Cross-references: GB:S72883; NID:G633864; PIDN:AAB31800.1; PID:G633865
R;Brook, J.D.; McCurrach, M.E.; Harley, H.G.; Buckler, A.J.; Church, D.; Aburatani, H.;
S.; Davies, J.; Shelbourne, P.; Buxton, J.; Jones, C.; Juvonen, V.; Johnson, K.; Harper,
Cell 68, 799-808, 1992
A;Title: Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repea
A;Reference number: A42101; MUID:92154692; PMID:1310900
A;Accession: A42101
A;Molecule type: mRNA
A;Residues: 'PPGQVRGRLLAVGGA', 57-550, 'WLWASRWGQA', 564-624 <BRO>
A;Cross-references: GB:M94203; GB:M91465; NID:G186755; PIDN:AAA64884.1; PID:G186756
A;Note: sequence extracted from NCBI backbone (NCBIN:82640, NCBI:P:82641)
C;Genetics:
A;Gene: GDB:DM
A;Cross-references: GDB:119097; OMIM:160900
A;Map position: 19q13.2-19q13.3
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; phosphotransferase
F;69-339/Domain: protein kinase homology <KIN>
F;77-85/Region: protein kinase ATP-binding motif

Query Match      8.4%; Score 878.5; DB 2; Length 624;
Best Local Similarity 37.5%; Pred. No. 6.9e-18;
Matches 194; Conservative 104; Mismatches 178; Indels 41; Gaps 11;

Qy      46 LSRREGILDALFVFEBCSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVSRLVGC 105
Db      20 LGLEPLDLLLGVHQLGASLQADKYVADFQWAPFIVVRLKEVRLQDRDDFEILKVGIR 79
Qy      106 GHFAVOVVRKATGDIYAMKMKKALLAQEOVSPFEERNILSRSTSPWITQLOAFQ 165
Db      80 GAFSEAVVRKQTGVYAMKMKKALLAQEOVSPFEERNILSRSTSPWITQLOAFQ 165
Db      80 GAFSEAVVRKQTGVYAMKMKKALLAQEOVSPFEERNILSRSTSPWITQLOAFQ 139
Qy      166 DKHLYLMEEYQPGDGLLSILNRYEDQDENLIQFYLAELILAVSHVLMGVVHRDIKPE 225

```

```

Db      140 DENLYLWMEYVVGDDLTLLSKFGERIPAEIMARFYLAETIWMADSVHRLGYVHRDIKPD 199
Qy      226 NILVDRTGHKILVDGSAAMKMNKNMNAKLPIGTDPYMAPEVL-TVMNGDGKGTGLDC 284
Db      200 NILDRGCHRLADDFGCLKLRADGTVRSILVAVGTPDYLSEILOAVGGGPGTGSVGP 259
Qy      285 DWMSGVGIAYEMIIYGRSPFAEGTSARTFNIMNFQRLKFP-DPPKVSSDFDLIOLSLC 343
Db      260 DWALGVFAVEMFYGTQPFVADSTAEYGIKVHYKHELSLPLVDEGVPEEARDFIQLLC 319
Qy      344 GQKERLXPEG---LCCHPFFSKIDMNNIRNSPPFVPTLKSDDDDTSNFDPEKSNVSS 400
Db      320 PPETRLGRGGAGDFRTHPFFFGDWDGLRDSVPFFPDFEGATDTCNFDLVEDGLTAMET 379
Qy      401 PCQLSPSGFSGEELPFVGFSGYKALGILGRSESVSGLDSPAKTSSMEKKL---INKE 457
Db      380 LSDIREGAPLGVLHLPFVGYSYS---CMALRDSEVPG---PTMEVEAEQLLEPHVQAPS 432
Qy      458 LODS---QDKCHKM-----EQEMTLHRRYSEVAVLSQKEV--ELKASETQRS 501
Db      433 LEPSVSPQDETAEVAVPAAPAAEAEBVLRLQELAELEBEVLTRSLRSEMAIRTD-- 490
Qy      502 LLEQLATYITECSSLKRSLE-----QARMEVSQED 532
Db      491 --NQNFASQLREAFARNEDLEAHVRQLQERMEQLQAE 525

RESULT 11
S71829
serine/threonine-specific protein kinase (EC 2.7.1.-) - mouse
N;Alternate names: myotonic dystrophy-associated protein kinase; myotonin protein kinase
C;Species: Mus musculus (house mouse)
C;Date: 06-Dec-1996 #sequence revision 06-Dec-1996 #text change 28-May-1999
C;Accession: S71829; S38815
R;Mahadevan, M.S.; Anemiy, C.; Jansen, G.; Sabourin, L.; Baird, S.; Neville, C.E.; Worme
submitted to the EMBL Data Library, January 1993
A;Description: Structure and genomic sequence of the myotonic dystrophy kinase (DM kinase
A;Reference number: S71829
A;Accession: S71829
A;Molecule type: DNA
A;Residues: 1-557 <MAH>
A;Cross-references: EMBL:Z21505
R;Mahadevan, M.S.; Anemiy, C.; Jansen, G.; Sabourin, L.; Baird, S.; Neville, C.E.; Worme
Hum. Mol. Genet. 2, 299-304, 1993
A;Title: Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene.
A;Reference number: S38815; MUID:93271990; PMID:8499920
A;Accession: S38815
A;Molecule type: DNA
A;Residues: 1-53 <MAW>
A;Cross-references: EMBL:Z21503
C;Genetics:
A;Introns: 54/1; 84/3; 112/3; 144/3; 194/2; 225/3; 294/3; 382/3; 411/2; 448/3; 503/2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; phosphotransferase; protein kinase
F;69-339/Domain: protein kinase homology <KIN>
F;77-85/Region: protein kinase ATP-binding motif

Query Match      8.2%; Score 862; DB 2; Length 557;
Best Local Similarity 36.6%; Pred. No. 1.8e-17;
Matches 199; Conservative 104; Mismatches 191; Indels 50; Gaps 13;

Qy      46 LSRREGILDALFVFEBCSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVSRLVGC 105
Db      20 LGLEPLDLLLGVHQLGASLQADKYVADFQWAPFIVVRLKEVRLQDRDDFEILKVGIR 79
Qy      106 GHFAVOVVRKATGDIYAMKMKKALLAQEOVSPFEERNILSRSTSPWITQLOAFQ 165
Db      80 GAFSEAVVRKQTGVYAMKMKKALLAQEOVSPFEERNILSRSTSPWITQLOAFQ 139
Qy      166 DKHLYLMEEYQPGDGLLSILNRYEDQDENLIQFYLAELILAVSHVLMGVVHRDIKPE 225
Db      140 DENLYLWMEYVVGDDLTLLSKFGERIPAEIMARFYLAETIWMADSVHRLGYVHRDIKPD 199

```


Db 1127 GESTKSIYQH---STLLELDQDKQAKYFLIPAMLNSEPA-----SRLMKRGKGLHIY 1177
Qy 1387 IPRFNVGLNRR-ATKCAVCLDTVH--FGRQASKLEQVMCHPKSTCLPATCGLPDABY 1443
Db 1178 NDHTF-VAVXVKGATCNVQQIRSFSSKQAYQCRDCRMVCHKTC-----1222
Qy 1444 ATHF-TEAFC 1452
Db 1223 --HYKTDAPC 1230

RESULT 13
G86431
protein kinase T518.9 protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 19-Jul-2002
C/Accession: G86431
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: G86431
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-522 <STO>
A/Cross-references: GB:AE005172; NID:g4587520; PIDN:AAD25751.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1
C/Supfamily: Arabidopsis thaliana protein kinase T518.9; protein kinase homology

Query Match 6.8%; Score 712; DB 2; Length 522;
Best Local Similarity 35.7%; Pred. No. 2.2e-13;
Matches 157; Conservative 82; Mismatches 133; Indels 68; Gaps 9;

Qy 69 KIKHVSFVRKYSDTTAELOELQPSAKDFEVRSLVCGHFAEVQVVRKATGDIYAMKVM 128
Db 93 KMDILKNFKENYK-M-RLQRQXGVDDFELLTWMIGKAFGEVRICKSTGVSFAMKGL 151
Qy 129 KKCALLAQEQVSPFEERNILSRSTSPWIPOLQYAFQDKXNHLYLMEYQPGDGLLSLNR 188
Db 152 KKSEMLRRGQVHVAERNVLAEDVSPFIVKLCYSQDDDEHLVLYMEYLPGGDMTLLMR 211
Qy 189 YEDOLDENLQFLAELLILAVSHVLMGVVHRDIKENILVDRTHGHIKLVDFGSAKWN 248
Db 212 -KUTLREDFRYVAQTILAIETSIKHNYIHRDIKPDNLLITRNGHIKLSDFGLSKLES 270
Qy 249 -----SNKVNNAKLPIGTPDYMAPEVLTVMNGDG 276
Db 271 KNFPDFKAEIVDRSTKPAAEHRLSKPPSPRRTQEQQLLHWQNRRTLAFSTVGTPTYI 330
Qy 265 APVLVVMGDKGTGLDCDHWKSVGIAYEMIVGRSPFAEGTSARTFNNIMNFORFLKFP 324
Db 331 APEVLV-----KKGVMGECMDWSLGAIMEFVLVGPFPFSEPLATCRKIYNWTKLKF 384
Qy 325 PDPKVSDFDLIQLCGQERLKFEG---LCCHPFFSKIDWNIRNSPPFPVPTLKS 381
Db 385 PDEAKLSIEVKLIRLLCNVQRLGTGKHVHKAPWFRGVEWELYESNAPYIPQVXH 444
Qy 382 DDTSNFDPEPKNSVSSPCQLSPSG-----FGSEELPFVGFYSKALGILGRSESV 435
Db 445 ELDTQNFEPDE-----VPSTCQTSSKSPWRKMSKDNANFLGTGYF-KNLEIV--DEHHI 497
Qy 436 SGLDSPAKTSSMEKLLIKS 455
Db 498 PGVAELKRKSKTANKPSLKT 517

RESULT 14

S42864
protein kinase (EC 2.7.1.-) - common ice plant (fragment)
C/Species: Mesembryanthemum crystallinum (common ice plant)
C/Date: 06-Jan-1995 #sequence_revision 23-Jul-1999 #text_change 19-Jul-2002
C/Accession: S42864
R/Baur, B.; Winter, K.; Fischer, K.; Dietz, K.
submitted to the EMBL Data Library, March 1994
A/Description: Molecular cloning and characterization of several protein kinases from pla
A/Reference number: S42864
A/Accession: S42864
A/Molecule type: mRNA
A/Residues: 1-479 <BAU>
A/Cross-references: EMBL:Z30329; NID:g457688; PIDN:CAA82990.1; PID:g457689
C/Supfamily: Arabidopsis thaliana protein kinase T518.9; protein kinase homology
C/Keywords: ATP; phosphotransferase; protein kinase
F:38-338/Domain: protein kinase homology <KIN>
F:46-54/Region: protein kinase ATP-binding motif

Query Match 6.7%; Score 706.5; DB 2; Length 479;
Best Local Similarity 35.1%; Pred. No. 2.9e-13;
Matches 163; Conservative 81; Mismatches 164; Indels 57; Gaps 8;

Qy 76 FVRKYSPTIAELQELQPSAKDFEVRSLVCGHFAEVQVVRKATGDIYAMKVMKKALLA 135
Db 19 FLEKETEYMRLORHKMGADDFELLTWMIGKAFGEVRVCRKTTGHVYAMKLLKXSEMLR 78
Qy 136 QEOVSFPEERNTLSRSTSPWIPOLQYAFQDKXNHLYLMEYQPGDGLLSLNRVEDOLDE 195
Db 79 RGQVHVKAERNLLAEVDNSNCIVKLYCSQDEEFLVHMEYLPFGDMTLLMR-KOTLIE 137
Qy 196 NLIQFYLAELILAVSHVLMGVVHRDIKENILVDRTHGHIKLVDFGSAKWN-----247
Db 138 DEARFYVAETILAIETSIKHNYIHRDIKPDNLLDKFGLHRLSDFGCLKPLDCLTEEKD 197
Qy 248 -----SNKVNNAKLPIGTPDYMAPEVLTVMNGDG 276
Db 198 FEVNGNGGSPNBSGSKPRRTQEQQLQHWQNRRLMAYSTVGTPTYIAPEVLL-----251
Qy 277 KGTGLDCDHWKSVGIAYEMIVGRSPFAEGTSARTFNNIMNFORFLKFPDPPKVSDFLD 336
Db 252 KKGVMGECMDWSLGAIMEFVLVGPFPFSDPSTCRKIYNWTHLKFPEEAKLSPEAKD 311
Qy 337 LIQSLCGQERLKFEG---LCCHPFFSKIDWNIRNSPPFPVPTLKSDDDTSNFDEPEK 393
Db 312 LISKLLCNVTQRLGSGNAHEIKLHPWFNGIDWERIYQMEAAFIPEVNDDELDTQNFKEE 371
Qy 394 NSWSSSPCOLSP--SGFGEELPFVGFYSKALGILGRSESVVSGLDSPAKTSSMEKLL 451
Db 372 ADNSQSQTSGAGPWRKMLSKDLNLFVGYTY-KNFEIV--NDYQVPGIAELKKDTPKRP 428
Qy 452 LIKS---KEIQDSQDKCHQWEQEMTRLHRRVSEVAVLSQKEVEL 493
Db 429 SIKSLFDESSDSSEATSGDQSVQSGFLNLLPQLLEVSTQTEV 473

RESULT 15

T26101
hypothetical protein W02B8.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T26101
R/Sims, M.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z20151
A/Accession: T26101
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1256 <WIL>
A/Cross-references: EMBL:Z81136; PIDN:CAB03458.1; GSPDB:GN00020; CESP:W02B8.2
A/Experimental source: clone W02B8

C;Genetics:
A;Gene: CFSF:W02B8.2
A;Map position: 2
A;Introns: 27/3; 327/3; 670/3; 949/3; 1073/3

Query Match 6.7%; Score 706.5; DB 2; Length 1256;
Best Local Similarity 21.1%; Pred. No. 7.1e-13;
Matches 308; Conservative 292; Mismatches 528; Indels 333; Gaps 54;

QY 545 QSKKLOIEKEQFYQAQVEENRMLNMNQLLEEDLVARSRRSDLYSELSERSLAA-EFVKKA 603
DB 10 QSSKTPKSKARYSSPYQVIDL--NNESQEDLTKRLLEAENIIQDLASERDALHESLVDKA 68
QY 604 TEOCHKLLKAKDOCKPEVGEYAKLEKINAEQQKIOELQKLSKAAKERAERELEKLQNR 663
DB 69 GLNESVLIERSNKVSTQETRIYRDEVTLLEDDLKQKESQIRIILQNRCLRLETEKQWQ-- 126
QY 664 EDSSEGIRKKLVBAEERRHSLNKKVRKLEFMERRENRLKDDIOTKSOOI-----QQVADKI 719
DB 127 -DTISGYQEDLKENE-----TRIENLSRLHLKLELSAKTHEIFSIGEELKNKT 175
QY 720 LELEEKHRE-----AOVSAOHLEVHLKQKQHYEKKIVLDNIOIKDLADKLETENMQR 774
DB 176 MKLNEKNSQOTKLAEISSNR--NLERKVKQFREELIVKD-----QRSLEV 220
QY 775 HESEAHKKGILSEQKAMINAMD--SKIRLEQRIVELSEANKLAANSUFTQRNKAQE 832
DB 221 HQDQENTQ-KVLKEVQLSDRLDYLTPKPKDVSRIKERDD-----FLOFSAKIIE 269
QY 833 EMISELRQQRKFXLEYETQAGKLEAQNRLKEGLEKISHQDHSDKRLLLETRLEVSLEHE 892
DB 270 ETWSLKLQVRLERELUSE-----KEELVKVT-----KEELQELQVTVQMGDSE 315
QY 893 E-----QKLELRQUTELQLSQERESQITLQAARAALLESQLOAKTELEETT--A 942
DB 316 QATKYLHAENMKLTROKADIRCDLLEARRNLKGFDEKRELEKQDEALEADVRRITELKK 375
QY 943 EASEETQALTAHDEIQKEDARNSC---TYVITLLEEQNLITEDNAELNNQNFLVK 998
DB 376 NVTEIURLSKLAEEREQIDEUKSRVAGYEVURDRHEAVYNLEAKAEKKNQGAHL-- 433
QY 999 QLDEASGANDEIVQLRSEVDHLRREITEREMQLTQKQTMALKTCTMLEEQMDLEAL 1058
DB 434 -----VMADKQSSHFK--TLKETAGSRRRAIE---QCNEVMVARIRGLEA- 473
QY 1059 NDELLKERQWEAWRSVLGDEKQFCQRCVRELQEMLDTEKQSRADORITERSQVVELA 1118
DB 474 ---SLENQRKVE-----QELE-MVKAENVROAKKLEFMKETEIQETHLD 512
QY 1119 VKSEKAEILALQALKEQKLKAESLSD-KLNDLEKHAMLEMMARSLOQKLETERELKQR 1177
DB 513 YREELSKL----AKGGSHEADSQSDSELSAKTTIQEVKADNKKLQILEEVQRQNSK 567
QY 1178 LLEBQAKLQOQMDLQKNHIFELTCQLOEALDRADLLKT-----ERSDLYQLENIQ 1228
DB 568 VLEENVKLR-----KGMABAELEKIEEFKRNWHSGREAGERLQLSAKENEK 613
QY 1229 VLYSHKVKVMKG-TISOQTKLIDFLQAKMQPAPKKKGLFSRRKEDPALPTQVPLQVNEL 1287
DB 614 VLKVEEELQEKRLVELEKEELVNYLQSQINTKQTKQPKL--SRSTLMSTISEV----DTS 668
QY 1288 KLALEKKAECABELEALQKTRIELRSARBEAAHRKATDHPHSTPATARQQATMSAIR 1347
DB 669 TYRVREVEVEEA--LEEQREELQVLAEBKRWADLQKSRSTANTTLLTS-----TT 717
QY 1348 SPEHOPSAMSLAPPSRRKSESPTPEFSRRLKERMEHNTPHRFNVGLNMR-----ATKC 1402
DB 718 ATEVSKSACELSYZPAT-----MRHDIPKW--LEFRHYSVLMSMKC 756
QY 1403 AVLCDTVHGRQASKCLCEQVMCHPKGKSTCLPATCGLPDAEVAHTFEAFORDKMSGLQ 1462
DB 757 SLCFVGTFTFTKANKGCHCDYRVHASCAPRVYNNTCGIPACQATYQDA--QAVGGIG-- 811

Qy	1463	TKPESSLLHLEGNMKVPRNNKRGQQQWDRKYIIVLEGSKVLIIYDNEAREACQRPVEE----	1511
Db	812	-----RNGWLRVYRD-LPESTWTSILWAMWDSNHVKFY-----RDAGADNLNPPYFT	858
Qy	1519	-----FELCLPDGDV-----SIHGAVGASELANATAKAKABADAK	1553
Db	859	IDLNKEQWILRTGOEVAIP-GDVMRNVLTIKLQTRSVH-----IVAPTPKS--ABRWAA	909
Qy	1554	LLGNS-----LLKLEGDDR-LDNACTLPFS-----DQVVLVGTGEEGYALNLV	1595
Db	910	CLQNAQTRRMKMSKPSAEYSCLLTLSLUPNNLKTFKAHTIEDWILFATQTLGFTFTSIS	969
Qy	1596	K-NSLTHVPIGIGAVFOYIYIKLEKLMTIAGEERALCLVDVKVKOSLAOSHPLPAQPDIS	1654
Db	970	QPRNPTRIAGPSSTVLEVMSEINCVMATNSKQLALPLDLSLTILAMQSTH-----PSIR	1025
Qy	1655	PNIFEAVKGCHLPGAGKIENGLCIICAMPKSVILRYNENLSKYCIKEIETSEPCSIH	1714
Db	1026	AEVLPEFGVHTIYHQDQDQREFLLVSDDTQLHRIKYNSTRDVFHAKVLVPEVPSTIE	1085
Qy	1715	FTNYSILIGNKFEYEDMKQYTLLEBFLDKNDHSLAPAFAS-----SNSFPVSIQVONS	1769
Db	1086	STAGIFASDTFYVYPL-----DHQSPSDVSARKLMPRRSDYVPVSAHITA	1133
Qy	1770	AGOREYLLCFHEFGVFDVSYGRRRTDDLKNSRLPLAFAYRPEPLYEVTHFNSLVEVIEQ	1829
Db	1134	-----NEILLAYQNGHIFVNLYGSGSRNQRIIEWEKPIEFAYTSPGLTYIVHDDSIIEVQIS	1189
Qy	1830	ARSSAGTPARAYLDIPNPRYLGPALSSGATYLAASSYQDKLVRVTCCKGNLVKESG-----	1883
Db	1190	K-----SSKEVTLA-----EREVIACVNAHIVQSDGVLSIVS	1221
Qy	1884	-----THHRGPSTSRSPNKR	1900
Db	1222	SKDITEVHRFESATCRGTGTR	1242

Search completed: July 3, 2004, 10:15:38
Job time : 44 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:17:35 ; Search time 105 Seconds
(without alignments)
6086.340 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10490

Sequence: 1 MLKFKYGARNPLDAGAEPI.....QLNGEIRQVEKSVLRTDYC 2053

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US50_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US50_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	10490	100.0	2053	13	US-10-017-216-2	Sequence 2, Appli
2	10490	100.0	2053	14	US-10-325-430-12	Sequence 12, Appl
3	10032.5	95.6	2054	12	US-10-415-011-21	Sequence 21, Appl
4	10022.5	95.5	2054	13	US-10-028-946-2	Sequence 2, Appli
5	10007	95.4	2066	12	US-09-964-956-9	Sequence 9, Appli
6	10007	95.4	2066	12	US-10-262-511-14	Sequence 14, Appl
7	10005	95.4	2053	12	US-09-964-956-11	Sequence 11, Appl
8	10005	95.4	2053	12	US-10-262-511-2	Sequence 2, Appli
9	9656	92.0	2055	13	US-10-017-216-4	Sequence 4, Appli
10	9487.5	90.4	1958	13	US-10-028-946-4	Sequence 4, Appli
11	7717.5	73.6	1841	12	US-09-964-956-40	Sequence 40, Appl
12	7717.5	73.6	1841	12	US-10-017-216-5	Sequence 5, Appli
13	7491.5	71.4	1597	12	US-09-964-956-41	Sequence 41, Appl
14	7491.5	71.4	1597	12	US-10-017-216-6	Sequence 6, Appli
15	6745.5	64.3	1441	15	US-10-412-897-3	Sequence 3, Appli

16	6223.5	59.3	1286	12	US-09-964-956-38	Sequence 38, Appl
17	6223.5	59.3	1286	13	US-10-017-216-7	Sequence 7, Appli
18	4718.5	45.0	999	12	US-10-276-774-1487	Sequence 1487, Ap
19	4519.5	43.1	940	12	US-09-964-956-39	Sequence 39, Appl
20	4249.5	40.5	883	12	US-09-964-956-37	Sequence 37, Appl
21	4101	39.1	847	12	US-10-262-511-8	Sequence 8, Appli
22	4005.5	38.2	832	12	US-10-262-511-6	Sequence 6, Appli
23	3155	30.1	623	12	US-10-262-511-4	Sequence 4, Appli
24	2910	27.7	623	12	US-10-262-511-10	Sequence 10, Appl
25	2892.5	27.6	638	12	US-10-262-511-12	Sequence 12, Appl
26	2430	23.2	497	9	US-09-804-471A-2	Sequence 2, Appli
27	2430	23.2	497	14	US-10-238-709-2	Sequence 2, Appli
28	2430	23.2	497	16	US-10-724-594-2	Sequence 2, Appli
29	2425	23.1	497	16	US-10-311-034-7	Sequence 7, Appli
30	2185.5	20.6	494	9	US-09-804-471A-4	Sequence 4, Appli
31	2185.5	20.6	494	14	US-10-238-709-4	Sequence 4, Appli
32	2185.5	20.6	494	16	US-10-724-594-4	Sequence 4, Appli
33	1700	16.2	373	15	US-10-120-835-42	Sequence 42, Appl
34	1645	15.7	319	15	US-10-412-897-2	Sequence 2, Appli
35	1446	13.8	1770	16	US-10-433-794-19	Sequence 19, Appl
36	1421.5	13.6	1719	12	US-10-362-892-2	Sequence 2, Appli
37	1421.5	13.6	1719	15	US-10-288-798-2	Sequence 2, Appli
38	1400	13.3	1738	15	US-10-210-130-100	Sequence 100, App
39	1399.5	13.3	1664	15	US-10-210-130-102	Sequence 102, App
40	1382.5	13.2	1711	9	US-09-771-161A-219	Sequence 219, App
41	1382.5	13.2	1711	9	US-09-771-161A-220	Sequence 220, App
42	1382.5	13.2	1711	12	US-10-399-225-10	Sequence 10, Appl
43	1277	12.2	257	14	US-10-282-048-2	Sequence 2, Appli
44	1258	12.0	1572	12	US-10-333-314-20	Sequence 20, Appl
45	1245	11.9	1548	15	US-10-369-493-6347	Sequence 6347, Ap

ALIGNMENTS

RESULT 1
US-10-017-216-2
; Sequence 2, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPILLER-LIBERMAN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote
; FILE OF INVENTION: Kinase and Uses Therefor
; TITLE OF INVENTION: Kinase and Uses Therefor
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2

Query Match	100.0%;	Score	10490;	DB	13;	Length	2053;	
Best Local Similarity	100.0%;	Pred. No.	0;					
Matches 2053;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MLKFKYGARNPLDAGAEPIASRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFE	60					
Db	1	MLKFKYGARNPLDAGAEPIASRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFE	60					
Qy	61	ECSPALMKIKHVSFNFKYSYSDTIAELQELQPSAKDFEVSRLVGGCGHFAEVQVVRKATG	120					
Db	61	ECSPALMKIKHVSFNFKYSYSDTIAELQELQPSAKDFEVSRLVGGCGHFAEVQVVRKATG	120					
Qy	121	DIYAMKMKKALLAQEQVSPFEERNILSRSTPMIPOLOYAFQDKNHLYLMEEYQGG	180					
Db	121	DIYAMKMKKALLAQEQVSPFEERNILSRSTPMIPOLOYAFQDKNHLYLMEEYQGG	180					
Qy	181	DLLSLNRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIPKPNILVDRGTGHIKLVDF	240					

181 DLSLNRVYEDQDENITQYLAELIIVAVSHVLMGYVHRDIKPENILVDRTHIILKLVDF 240
241 GSAAKNNSNMVNAKLPIGTDFDYMAPEVLTVMNGDGKTYGLDCCDWMSGVIAIYGR 300
241 GSAAKNNSNMVNAKLPIGTDFDYMAPEVLTVMNGDGKTYGLDCCDWMSGVIAIYGR 300
301 SPFAEGTSARTFNINMFQRLKFPDPPKVSDFDLIOQLLGCQKRLKFBGLCCHPFF 360
301 SPFAEGTSARTFNINMFQRLKFPDPPKVSDFDLIOQLLGCQKRLKFBGLCCHPFF 360
361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPKNSWVSSPCOLSPSGFSGBELPVGFS 420
361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPKNSWVSSPCOLSPSGFSGBELPVGFS 420
421 YSKALGILGRSESVVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHMEQEMTRLHRRVS 480
421 YSKALGILGRSESVVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHMEQEMTRLHRRVS 480
481 EYSAVLSQKVELKASRTQSLAEQDLATYITECSSIKRSLEQARMEVSQEDDKALQLLH 540
481 EYSAVLSQKVELKASRTQSLAEQDLATYITECSSIKRSLEQARMEVSQEDDKALQLLH 540
541 DIREQSKLOEIKQEYQAOVEEMRLMNNQLEEDLVSAARRSDLYESELRESLAAEFK 600
541 DIREQSKLOEIKQEYQAOVEEMRLMNNQLEEDLVSAARRSDLYESELRESLAAEFK 600
601 RKATECOHLLKAKOQKPVGYAKLEKINAEQQLKIOELOKLEKAKERAERELEKL 660
601 RKATECOHLLKAKOQKPVGYAKLEKINAEQQLKIOELOKLEKAKERAERELEKL 660
661 QNRDSSEGIRKKLVEAEERHSLNKVKLEMTERRNRLKDDIOTKSSQQIQQVADKIL 720
661 QNRDSSEGIRKKLVEAEERHSLNKVKLEMTERRNRLKDDIOTKSSQQIQQVADKIL 720
721 ELBEKHREAQVSAQHLVHLKQKQHYEYKIKVLDNOIKDLADKETLENMMQHEEBAH 780
721 ELBEKHREAQVSAQHLVHLKQKQHYEYKIKVLDNOIKDLADKETLENMMQHEEBAH 780
781 EKGKILSEQAMNAMDKIRSLRQIVELSEANKLAANSSLTQRMNKAQEMISELRQ 840
781 EKGKILSEQAMNAMDKIRSLRQIVELSEANKLAANSSLTQRMNKAQEMISELRQ 840
841 QKPYLETQAGKLEAQNRLKEQLEKISHQDHSQKRLLELETLRLRVSLEHEEQKLELR 900
841 QKPYLETQAGKLEAQNRLKEQLEKISHQDHSQKRLLELETLRLRVSLEHEEQKLELR 900
901 QLTLEQLSQRSELTALQARAALLESQRLQAKTELETTAEABEEIOALTATHRDEIOR 960
901 QLTLEQLSQRSELTALQARAALLESQRLQAKTELETTAEABEEIOALTATHRDEIOR 960
961 KFDALRNSCTVITDLEBQLNQLTEDNAELNNQNFYLSKQIDEASGANDEIVQLRSEVDHL 1020
961 KFDALRNSCTVITDLEBQLNQLTEDNAELNNQNFYLSKQIDEASGANDEIVQLRSEVDHL 1020
1021 RREITEREMQITQKQTMELKTTCTMLEBQVNDLEALDELLEKESQWENRSEVLGDEK 1080
1021 RREITEREMQITQKQTMELKTTCTMLEBQVNDLEALDELLEKESQWENRSEVLGDEK 1080
1081 SQECPVREIQLRMLDTEKQSRARADQRTESROWELAVKEHKAETILAQALKEQKLA 1140
1081 SQECPVREIQLRMLDTEKQSRARADQRTESROWELAVKEHKAETILAQALKEQKLA 1140
1141 ESLSKLNLEKHAMLENMARSLOQKLETERELKORLLEBQAKLOQMDLQXNHFRLT 1200
1141 ESLSKLNLEKHAMLENMARSLOQKLETERELKORLLEBQAKLOQMDLQXNHFRLT 1200
1201 QGQEAALDRADLLKTERSDEYOLENIQVLYSHEKVMGTTISQQTKLDFLOAKMDQPA 1260
1201 QGQEAALDRADLLKTERSDEYOLENIQVLYSHEKVMGTTISQQTKLDFLOAKMDQPA 1260
1261 KKKKGLFSRRKEDPALPTQVPLQYNELKLALEKEKARCABELEALQKTRIELRSAREBA 1320

1261 KKKKGLFSRRKEDPALPTQVPLQYNELKLALEKEKARCABELEALQKTRIELRSAREBA 1320
1321 HRKATDHPSTPATARQQAISAIVRSPHQPSPMSLLAPPSSRRKESSTPPEFSRLK 1380
1321 HRKATDHPSTPATARQQAISAIVRSPHQPSPMSLLAPPSSRRKESSTPPEFSRLK 1380
1381 BRMHNTPHRENVLNMRATKCAVCLDTVHFGRAQSKLEQVCMCHPKCSTCIPATCGLP 1440
1381 BRMHNTPHRENVLNMRATKCAVCLDTVHFGRAQSKLEQVCMCHPKCSTCIPATCGLP 1440
1441 AEYATHTEAFCRDKMNSPGLQTKPESSSLHLEBGMKVPRNNKRGQOGWDRKIVILEGSK 1500
1441 AEYATHTEAFCRDKMNSPGLQTKPESSSLHLEBGMKVPRNNKRGQOGWDRKIVILEGSK 1500
1501 VLIYDNEAREAGQPVVEEFELCLPDGDSVTHGAVGASELANATAKAEKADAKLGNLSL 1560
1501 VLIYDNEAREAGQPVVEEFELCLPDGDSVTHGAVGASELANATAKAEKADAKLGNLSL 1560
1561 KLEGDDRLDMNCTLPFSDQVVLVGTBGLYALNVLKNSLTHVPGIGAVFQIYIKDLEKL 1620
1561 KLEGDDRLDMNCTLPFSDQVVLVGTBGLYALNVLKNSLTHVPGIGAVFQIYIKDLEKL 1620
1621 LMIAGEERALCLVDVKVQSLAQSHLPAQPDISPNTFEAVKGCCHLFGACKIENGLCICA 1680
1621 LMIAGEERALCLVDVKVQSLAQSHLPAQPDISPNTFEAVKGCCHLFGACKIENGLCICA 1680
1681 AMPKSVILRYNENLSKYCIRKEIETSEPCSIHFTNYSILIGTNKFEYIDMKQYTLLEEF 1740
1681 AMPKSVILRYNENLSKYCIRKEIETSEPCSIHFTNYSILIGTNKFEYIDMKQYTLLEEF 1740
1741 LDKNDHSLAPAVFAASNSPPVSIQVNSAQOREYLLCFHEFGVFDVSYGRSRTDCLK 1800
1741 LDKNDHSLAPAVFAASNSPPVSIQVNSAQOREYLLCFHEFGVFDVSYGRSRTDCLK 1800
1801 WSRPLAPAYREPVLVTHFNLSLEVEIQARSSAGTTPARAYLDIPNRYLGPALISSGAIY 1860
1801 WSRPLAPAYREPVLVTHFNLSLEVEIQARSSAGTTPARAYLDIPNRYLGPALISSGAIY 1860
1861 LASSYQDKRLVICCKGNLVKESGTEHHRGPTSSRSPNKGPPTYNETHITKRVASSPAPP 1920
1861 LASSYQDKRLVICCKGNLVKESGTEHHRGPTSSRSPNKGPPTYNETHITKRVASSPAPP 1920
1921 EGPSPREPSTPHRYRGRTRELREDKSPGRPLEREKSPGRMLSTRRERSPGRLPEDSSRG 1980
1921 EGPSPREPSTPHRYRGRTRELREDKSPGRPLEREKSPGRMLSTRRERSPGRLPEDSSRG 1980
1981 RLPAGAVRTPLSQVNSKRGQSASQVFTVNTVYVYDMNKKLDNLPANWSVLRIIQLNGEIR 2040
1981 RLPAGAVRTPLSQVNSKRGQSASQVFTVNTVYVYDMNKKLDNLPANWSVLRIIQLNGEIR 2040
2041 QQVEKSVLRDTC 2053
2041 QQVEKSVLRDTC 2053

RESULT 2

US-10-325-430-12
; Sequence 12, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MP101-294P1RNM
; CURRENT APPLICATION NUMBER: US/10/325.430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 12
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-325-430-12

Query Match      100.0%; Score 10490; DB 14; Length 2053;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPIASGRSLNLFPGGKPPFTTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFKYGARNPLDAGAAEPIASGRSLNLFPGGKPPFTTQQQMSPLSREGILDALFVLFE 60

QY 61 EGSQPALMKIKHVSFNFRKYSYDTIAELOQLOPSAKDFEVRSLVGCCHFAEYQVVRKATG 120
Db 61 EGSQPALMKIKHVSFNFRKYSYDTIAELOQLOPSAKDFEVRSLVGCCHFAEYQVVRKATG 120

QY 121 DIYAMKVMKKALLAQEQVSFFEBERNILSRSTSPWIPOLQYAFQDKNHLVMEBYQPGG 180
Db 121 DIYAMKVMKKALLAQEQVSFFEBERNILSRSTSPWIPOLQYAFQDKNHLVMEBYQPGG 180

QY 181 DLLSLNRYVEDQDENLTOFYLAELILAVHSVHLGYVHRDIKPNILVDRTHGHIKLVDF 240
Db 181 DLLSLNRYVEDQDENLTOFYLAELILAVHSVHLGYVHRDIKPNILVDRTHGHIKLVDF 240

QY 241 GSAAKWNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGYGLDCDWSVGVIAYEMIYGR 300
Db 241 GSAAKWNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGYGLDCDWSVGVIAYEMIYGR 300

QY 301 SPFAGTSARTNNIMNFORFLKFPDDPKVSDDFLDLQSLCGOKERLKEPGLCCHPFF 360
Db 301 SPFAGTSARTNNIMNFORFLKFPDDPKVSDDFLDLQSLCGOKERLKEPGLCCHPFF 360

QY 361 SKIDWNNIRNSPPFPVPTLKSDDTSNFDPEPKNSWSSPCQLSPGSGEELPFVGFS 420
Db 361 SKIDWNNIRNSPPFPVPTLKSDDTSNFDPEPKNSWSSPCQLSPGSGEELPFVGFS 420

QY 421 YSKALGILGRSSVSGLDSPAKTSMEKKLIKSELQDSQDKCHQMEQETRLHRRVS 480
Db 421 YSKALGILGRSSVSGLDSPAKTSMEKKLIKSELQDSQDKCHQMEQETRLHRRVS 480

QY 481 EYEAIVLSQKEVELKASETORSLEODLATYITECSSLSRSLQEARMEVSQEDDKALQLLH 540
Db 481 EYEAIVLSQKEVELKASETORSLEODLATYITECSSLSRSLQEARMEVSQEDDKALQLLH 540

QY 541 DIREQSRKLQETKEQEYQAEVEMRLMNQLEEDILVSARRSDLYESELRESLAAEEFK 600
Db 541 DIREQSRKLQETKEQEYQAEVEMRLMNQLEEDILVSARRSDLYESELRESLAAEEFK 600

QY 601 RKATCOHKLILAKQCGKPEVGEYAKLEKINAEQOLKIQELQEKLEKAAKRAERLEBKL 660
Db 601 RKATCOHKLILAKQCGKPEVGEYAKLEKINAEQOLKIQELQEKLEKAAKRAERLEBKL 660

QY 661 QNREDSSEIGIRKKLVEAEBRRHSLENKVKRLTMRERENRLKDDIQTKSQOIQOWADKIL 720
Db 661 QNREDSSEIGIRKKLVEAEBRRHSLENKVKRLTMRERENRLKDDIQTKSQOIQOWADKIL 720

QY 721 ELEKHREAAQVSAQHLEVLKQKEHYBEKIKVLQNKQKDLADKETLENVQVREHEBAH 780
Db 721 ELEKHREAAQVSAQHLEVLKQKEHYBEKIKVLQNKQKDLADKETLENVQVREHEBAH 780

QY 781 EKGKILSEOKAMINAMDSKIRSLQRIVELSEANKLAANSILFTORNKKAQEEIMSELRO 840
Db 781 EKGKILSEOKAMINAMDSKIRSLQRIVELSEANKLAANSILFTORNKKAQEEIMSELRO 840

QY 841 QKPYLETQAGKLEAQNRLKEOLEKISHQDSDKNRLLELETRLRREVSLHEEQKLELR 900
Db 841 QKPYLETQAGKLEAQNRLKEOLEKISHQDSDKNRLLELETRLRREVSLHEEQKLELR 900

QY 901 QUTELQLSQERESQITALQAARAALESQOLROAKTELETTAAABEEIQALTARDEIQR 960
Db 901 QUTELQLSQERESQITALQAARAALESQOLROAKTELETTAAABEEIQALTARDEIQR 960
```

```

961 KFDALRNSCTVITDLBEQLNQLTEDNAELNQNPFYLSKQLDREASGANDEIVQIRSEVDHL 1020
961 KFDALRNSCTVITDLBEQLNQLTEDNAELNQNPFYLSKQLDREASGANDEIVQIRSEVDHL 1020

1021 RREITEREMQLTQKOTMEALKTTCTMLBEQVMDLEALNDELLEKEREQWEARSWVLGDEK 1080
1021 RREITEREMQLTQKOTMEALKTTCTMLBEQVMDLEALNDELLEKEREQWEARSWVLGDEK 1080

1081 SQFECRVRELQRLMDTEKQSRARADQRIETESQVVELAVKHEKAEILALQOALKEQKLKA 1140
1081 SQFECRVRELQRLMDTEKQSRARADQRIETESQVVELAVKHEKAEILALQOALKEQKLKA 1140

1141 ESI5DKLNDLEKXHAMLENNARSILOQKLETERELKORLFEQAKLQOQMDLQKNHIFRLT 1200
1141 ESI5DKLNDLEKXHAMLENNARSILOQKLETERELKORLFEQAKLQOQMDLQKNHIFRLT 1200

1201 QGLQEALDRADLLKTERSDLEYOLENIQVLYSHEKVMGCTISQOTKILIDFLQAKMDQPA 1260
1201 QGLQEALDRADLLKTERSDLEYOLENIQVLYSHEKVMGCTISQOTKILIDFLQAKMDQPA 1260

1261 KKKKGLFSRKEOPALPTQVPLQVNEUKLALEKEKARCAELEALOKTRTELRSAREEA 1320
1261 KKKKGLFSRKEOPALPTQVPLQVNEUKLALEKEKARCAELEALOKTRTELRSAREEA 1320

1321 HRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSSRRKESSTPEEF5RRLK 1380
1321 HRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSSRRKESSTPEEF5RRLK 1380

1381 ERMHNTPHFENGLNMRATKCAVCLDTVHFGRQASKCLEQVCHPKCSTCIPATCGLP 1440
1381 ERMHNTPHFENGLNMRATKCAVCLDTVHFGRQASKCLEQVCHPKCSTCIPATCGLP 1440

1441 AEYATHTFEAFCDKXNSPGLQTKESPSSSLHLEGWVKVPRNKRGOQGWDRKVIIVLEGSK 1500
1441 AEYATHTFEAFCDKXNSPGLQTKESPSSSLHLEGWVKVPRNKRGOQGWDRKVIIVLEGSK 1500

1501 VLIYDNEARAGORPVEEFELCLPDGQVSIHGAVGASELANATAKAEKAEADAKLGN5LL 1560
1501 VLIYDNEARAGORPVEEFELCLPDGQVSIHGAVGASELANATAKAEKAEADAKLGN5LL 1560

1561 KLEGDDRLDNVCTLPFSDQVVLVGTBEGYALNVLKNSLTHVPGIGAVFOIYIIKOLEKL 1620
1561 KLEGDDRLDNVCTLPFSDQVVLVGTBEGYALNVLKNSLTHVPGIGAVFOIYIIKOLEKL 1620

1621 LMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCHLFGAGKIE5GLCICA 1680
1621 LMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCHLFGAGKIE5GLCICA 1680

1681 AMP5KVVILRYNENLSKYCIRKEIETSEPC5CTHFTNY5LITGNTKPYEIDMKQYTL5EFP 1740
1681 AMP5KVVILRYNENLSKYCIRKEIETSEPC5CTHFTNY5LITGNTKPYEIDMKQYTL5EFP 1740

1741 LDKNDSLAPAVFAA5SNSFPV5IVQVNSAGQRE5YLLCFHEFGVFVDSYGR5R5TDD5LK 1800
1741 LDKNDSLAPAVFAA5SNSFPV5IVQVNSAGQRE5YLLCFHEFGVFVDSYGR5R5TDD5LK 1800

1801 W5RLPLAFAYREPYLFVTHFNSLE5VIEIQAR5SAGTPARAYLDIPNPRYLGP5AIS5GAY 1860
1801 W5RLPLAFAYREPYLFVTHFNSLE5VIEIQAR5SAGTPARAYLDIPNPRYLGP5AIS5GAY 1860

1861 LASSYQDKLRVICCKGNLVKESGTEHHRG5P5TSR5SP5KRG5P5TYNEHITTKRV55SP5APP 1920
1861 LASSYQDKLRVICCKGNLVKESGTEHHRG5P5TSR5SP5KRG5P5TYNEHITTKRV55SP5APP 1920

1921 EGF5SHPRE5P5THRYREGRT5EL5RDK5PGK5PL5EREK5PGK5ML5TR5R5SP5GR5L5F5D5SR5G 1980
1921 EGF5SHPRE5P5THRYREGRT5EL5RDK5PGK5PL5EREK5PGK5ML5TR5R5SP5GR5L5F5D5SR5G 1980

1981 RL5FAGAVRTPL5QV5NKG5R5Q5ASQV5T5VNT5VY5DN5K5LD5NL5PAN5V5L5RI5QL5NG5EIR 2040
1981 RL5FAGAVRTPL5QV5NKG5R5Q5ASQV5T5VNT5VY5DN5K5LD5NL5PAN5V5L5RI5QL5NG5EIR 2040
```


QY	2041	QVQKSVLRDTC	2053	1	MLKFKYCARNPLDAGAEPIASRASRLNLFPOGKPPMTQOQMSPLSREGILDALFVLPE	60
DB	2041	QVQKSVLRDTC	2053	1	MLKFKYCARNPLDAGAEPIASRASRLNLFPOGKPPMTQOQMSPLSREGILDALFVLPE	60
RESULT 3						
US-10-415-011-21						
; Sequence 21, Application US/10415011						
; Publication No. US20040053394A1						
; GENERAL INFORMATION:						
; APPLICANT: INCYTE CORPORATION						
; APPLICANT: GURUPAJAN, Rajagopal						
; APPLICANT: BAUGHN, Mariah R.						
; APPLICANT: CHAWLA, Narinder K.						
; APPLICANT: ELLIOTT, Vicki S.						
; APPLICANT: XU, Yuming						
; APPLICANT: ARVIZU, Chandra S.						
; APPLICANT: YAO, Monique G.						
; APPLICANT: RAMKUMAR, Jayalaxmi						
; APPLICANT: DING, Li						
; APPLICANT: TANG, Yi Tom						
; APPLICANT: HAFALIA, April J.A.						
; APPLICANT: NGUYEN, Dannel B.						
; APPLICANT: GANDHI, Ameena R.						
; APPLICANT: LU, Yan						
; APPLICANT: YUE, Henry						
; APPLICANT: BURFORD, Neil						
; APPLICANT: BANDMAN, Olga						
; APPLICANT: TRIBOULEY, Catherine M.						
; APPLICANT: LAL, Preeti G.						
; APPLICANT: RECIPON, Shirley A.						
; APPLICANT: LU, Dyung Aina M.						
; APPLICANT: BOROMSKY, Mark L.						
; APPLICANT: THORNTON, Michael B.						
; APPLICANT: SWANAKER, Anita						
; APPLICANT: THANGAVELU, Kavitha						
; APPLICANT: KHAN, Farrah A.						
; APPLICANT: ISON, Craig H.						
; TITLE OF INVENTION: HUMAN KINASES						
; FILE REFERENCE: PI-0262 USN						
; CURRENT APPLICATION NUMBER: US/10/415,011						
; CURRENT FILING DATE: 2003-04-18						
; PRIOR APPLICATION NUMBER: PCT/US01/47728						
; PRIOR FILING DATE: 2001-10-20						
; PRIOR APPLICATION NUMBER: US 60/242,410						
; PRIOR FILING DATE: 2000-10-20						
; PRIOR APPLICATION NUMBER: US 60/244,068						
; PRIOR FILING DATE: 2000-10-27						
; PRIOR APPLICATION NUMBER: US 60/245,708						
; PRIOR FILING DATE: 2000-11-03						
; PRIOR APPLICATION NUMBER: US 60/247,672						
; PRIOR FILING DATE: 2000-11-09						
; PRIOR APPLICATION NUMBER: US 60/249,565						
; PRIOR FILING DATE: 2000-11-16						
; PRIOR APPLICATION NUMBER: US 60/252,730						
; PRIOR FILING DATE: 2000-11-22						
; PRIOR APPLICATION NUMBER: US 60/250,807						
; PRIOR FILING DATE: 2000-12-01						
; NUMBER OF SEQ ID NOS: 44						
; SOFTWARE: PERL Program						
; SEQ ID NO 21						
; LENGTH: 2054						
; TYPE: PRT						
; ORGANISM: Homo sapiens						
; FEATURE:						
; NAME/KEY: misc feature						
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7484498CD1						
US-10-415-011-21						
Query Match 95.6%; Score 10032.5; DB 12; Length 2054;						
Best Local Similarity 95.8%; Pred. No. 0;						
Matches 1981; Conservative						

Db 1081 KRFQEWASVLDGKESQFECVRRLQRMLDTEKQSRAPADQRITTESQVVELAYKEHA 1140
QY 1125 EIALAQALKEQKLAESLSDKLNLEKHAMLENNARSLOQKLTTERELKORLEBOAK 1184
Db 1141 EIALAQALKEQKLAESLSDKLNLEKHAMLENNARSLOQKLTTERELKORLEBOAK 1200
QY 1185 LOQMDLQXNHI,FRITQGOEALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1244
Db 1201 LOQMDLQXNHI,FRITQGOEALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1260
QY 1245 QTKLIDFLOAKQDQPAKKKGLFSRRKEDPALPTQVPOYNELKLALEKEKARCAELBEA 1304
Db 1261 QTKLIDFLOAKQDQPAKKK-----VPLQYNELKLALEKEKARCAELBEA 1305
QY 1305 LOKTRIELRSAREEAHRAKATDHPHPSTPATARQCIAMSAIVRSPEHOPSAWSLLAPSS 1364
Db 1306 LOKTRIELRSAREEAHRAKATDHPHPSTPATARQCIAMSAIVRSPEHOPSAWSLLAPSS 1365
QY 1365 RRKESSTPEFSRRLLKERMHHNIPHRFNVGLNMRATKCAVCLDTHVFGQSKCLECQVM 1424
Db 1366 RRKESSTPEFSRRLLKERMHHNIPHRFNVGLNMRATKCAVCLDTHVFGQSKCLECQVM 1425
QY 1425 CHPKSTCLPATCGUPAEYATHTFAFCDDKXNSPGLQTKPSSSIHLHLEGNMKVPRNNKR 1484
Db 1426 CHPKSTCLPATCGUPAEYATHTFAFCDDKXNSPGLQTKPSSSIHLHLEGNMKVPRNNKR 1485
QY 1485 GQGGWDRKVIYLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544
Db 1486 GQGGWDRKVIYLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1545
QY 1545 A-----EKARADAKL 1554
Db 1546 ADVPYILKMHSPHTTCWFGRTLYLLAPFPDKQRWTALESVWAGVRSREKAPADAKL 1605
QY 1555 LGNSLLKLEGGDRLDNCTLPFSDQVLVGTTEGLYALNVLKNSLTHVPGIGAVFQIYII 1614
Db 1606 LGNSLLKLEGGDRLDNCTLPFSDQVLVGTTEGLYALNVLKNSLTHVPGIGAVFQIYII 1665
QY 1615 KOLEKLLMAGBERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN 1674
Db 1666 KOLEKLLMAGBERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN 1725
QY 1675 GLCICAAKPSKVILRYNENLSKYCIKRIEITSEPCSCIHFNTYISILGTNKFYEDMKQ 1734
Db 1726 GLCICAAKPSKVILRYNENLSKYCIKRIEITSEPCSCIHFNTYISILGTNKFYEDMKQ 1785
QY 1735 YTLBFLDKNDHSLAPAVFAASNSFPVSIQVNSAGQREXYLLCFHFHFGVFDVSYGRRS 1794
Db 1786 YTLBFLDKNDHSLAPAVFAASNSFPVSIQVNSAGQREXYLLCFHFHFGVFDVSYGRRS 1845
QY 1795 RTDDLKWSRLPLAFAYREPYLFTVTHFNSLEVIEIQARSAGTPPARAYLDIPNPRYLGPAI 1854
Db 1846 RTDDLKWSRLPLAFAYREPYLFTVTHFNSLEVIEIQARSAGTPPARAYLDIPNPRYLGPAI 1905
QY 1855 SSGAIVLASSYQDKLRVICCKNLVKESGTEHHRGPSTSRSSPNKRGPPTTNEHITKRA 1914
Db 1906 SSGAIVLASSYQDKLRVICCKNLVKESGTEHHRGPSTSRSSPNKRGPPTTNEHITKRA 1965
QY 1915 SSPAPPEGSPHREBPTPHRYREGRTLRDCKSPGRPLEREKSPGRMLSTRERSPGRLF 1974
Db 1966 SSPAPPEGSPHREBPTPHRYREGRTLRDCKSPGRPLEREKSPGRMLSTRERSPGRLF 2025
QY 1975 EDSSRGRLPAGAVRTPLSQVKNKGQSA 2002
Db 2026 EDSSRGRLPAGAVRTPLSQVKNKWDQSS 2053

RESULT 4
US-10-028-946-2
; Sequence 2, Application US/10028946
; Publication No. US20020123622A1
; GENERAL INFORMATION:

; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2054
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-2

Query March 95.5%; Score 10022.5; DB 13; Length 2054;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;
QY 1 MLKFKYGARPLDAGAAEPIASRASRLNLFQGGPPMTQOQMSPLSREGILDALFVLF 60
Db 1 MLKFKYGARPLDAGAAEPIASRASRLNLFQGGPPMTQOQMSPLSREGILDALFVLF 60
QY 61 ECSQPALMKIKHVSFNVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVRKATG 120
Db 61 ECSQPALMKIKHVSFNVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVRKATG 120
QY 121 DIYAMKVMKKALLAQOVSFPEERNLRSRSPWLPOLQYAFQDKNHLVLMEEYOPGG 180
Db 121 DIYAMKVMKKALLAQOVSFPEERNLRSRSPWLPOLQYAFQDKNHLVLMEEYOPGG 180
QY 181 DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGVVHRDIPENILVDRTHGILVDF 240
Db 181 DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGVVHRDIPENILVDRTHGILVDF 240
QY 241 GSAAKMNSKNNAKLIBITPDYVAPVLTVMGDDGKGTGLDCDWMVSGVIAYEMIYGR 300
Db 241 GSAAKMNSKNNAKLIBITPDYVAPVLTVMGDDGKGTGLDCDWMVSGVIAYEMIYGR 300
QY 301 SPFAEGTSARTFNINMNFQFLKFPDPPKYSDFDLIQSLLCGQKRLKFEGLCCHPFF 360
Db 301 SPFAEGTSARTFNINMNFQFLKFPDPPKYSDFDLIQSLLCGQKRLKFEGLCCHPFF 360
QY 361 SKIDMNNIRNSPPFPVTLKSDDDTSNFDPEKNSWYSSSPCQLSPSGFSGEELPFVGF 420
Db 361 SKIDMNNIRNSPPFPVTLKSDDDTSNFDPEKNSWYSSSPCQLSPSGFSGEELPFVGF 420
QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLILKSKELQSDCKHKEQMTLHRVS 480
Db 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLILKSKELQSDCKHKEQMTLHRVS 480
QY 481 EVEAVLSQKEVELKASQTQSLLEQDLATYITECSSLKRSLKARMEVSDDDKALQLH 540
Db 481 EVEAVLSQKEVELKASQTQSLLEQDLATYITECSSLKRSLKARMEVSDDDKALQLH 540
QY 541 DIRQSKLQEIKEQYQAOVEEVRMLMNOLEBDLVSAARRSDLYSELSRLAAEFK 600
Db 541 DIRQSKLQEIKEQYQAOVEEVRMLMNOLEBDLVSAARRSDLYSELSRLAAEFK 600
QY 601 RKATCCQKLLKAKDQKPEVGEYAKLEKINAEQOLKIQELQEKLEA----- 648
Db 601 RKATCCQKLLKAKDQKPEVGEYAKLEKINAEQOLKIQELQEKLEKAVKASTEATELLQ 660
QY 649 -----AKERARELEKLNREDSDSEGIRKKLVAEERPHSLKVKRLTETMERENRLKDD 704
Db 661 NIROAKERARELEKLNREDSDSEGIRKKLVAEERPHSLKVKRLTETMERENRLKDD 720
QY 705 IOTKSOQIQOMADKILEEKKHREAOVSACHLEVHLKQKHOHYEELKVLJDNQIKKDLAD 764
Db 721 IOTKSOQIQOMADKILEEKKHREAOVSACHLEVHLKQKHOHYEELKVLJDNQIKKDLAD 780

QY 765 KETLENMQRHEBEAHEKGIILSEQAMINAMDSKIRSLERIVELSEANKLAANSSLFT 824
Db 781 KETLENMQRHEBEAHEKGIILSEQAMINAMDSKIRSLERIVELSEANKLAANSSLFT 840
QY 825 QRNKAQBEIMISBLROOKFYLETQAGKLEAQRNKLBEQLEKISHQDSKQNRILLELETRL 884
Db 841 QRNKAQBEIMI SELROOKFYLETQAGKLEAQRNKLBEQLEKISHQDSKQNRILLELETRL 900
QY 885 REVSLEHEBEOKLEIKQLTBLQSLQERESQLTALQAAARAALESQURQAkteLEETTABA 944
Db 901 REVSLEHEBEOKLEIKQLTBLQSLQERESQLTALQAAARAALESQURQAkteLEETTABA 960
QY 945 EEEIQALTARHDEIORKFDALRNSCTVITDLEBOLNQLTDENAELNNQNFYLSKOLDEAS 1004
Db 961 EEEIQALTARHDEIORKFDALRNSCTVITDLEBOLNQLTDENAELNNQNFYLSKOLDEAS 1020
QY 1005 GANDEIVQLRSEVDHLRREITEREMOLTQOKTMEALKTTCTMLEBQVMDLBALANDELLE 1064
Db 1021 GANDEIVQLRSEVDHLRREITEREMOLTQOKTMEALKTTCTMLEBQVMDLBALANDELLE 1080
QY 1065 KERQWEAARSVLGDSEKQRECRVRELQRMLDTEKQSPARADQRIETESROVVELAVKEHKA 1124
Db 1081 KERQWEAARSVLGDSEKQRECRVRELQRMLDTEKQSPARADQRIETESROVVELAVKEHKA 1140
QY 1125 EIIALQOALKEQKLAESLSDKLNDELKKHAMLENNARSLOQKLETERELKORLLEEQAQ 1184
Db 1141 EIIALQOALKEQKLAESLSDKLNDELKKHAMLENNARSLOQKLETERELKORLLEEQAQ 1200
QY 1185 LOQOMDLQKNHIFRLTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVKVMEGTISQ 1244
Db 1201 LOQOMDLQKNHIFRLTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVKVMEGTISQ 1260
QY 1245 QTKLIDFLOAKMDQPAKKKKGLFSRRKEDPALPTQVPLQYNELKLALKEKARCABLESEA 1304
Db 1261 QTKLIDFLOAKMDQPAKKKKGLFSRRKEDPALPTQVPLQYNELKLALKEKARCABLESEA 1305
QY 1305 LQKTRIELSARBEAARXATDHPHSTPATARQOIAMSAIVRSBPBOPSAMSLAPSS 1364
Db 1306 LQKTRIELSARBEAARXATDHPHSTPATARQOIAMSAIVRSBPBOPSAMSLAPSS 1365
QY 1365 RRKESSTPEEFSRRLKERVHNNIHPHFNVLNMRATKCAVCLDTVHFGRQASKCLECQVM 1424
Db 1366 RRKESSTPEEFSRRLKERVHNNIHPHFNVLNMRATKCAVCLDTVHFGRQASKCLECQVM 1425
QY 1425 CHPKSTCLPATCGLPAEYATHTFEACRDKNNSPGLQTKPESSSLHLEGMMKVPNNKR 1484
Db 1426 CHPKSTCLPATCGLPAEYATHTFEACRDKNNSPGLQTKPESSSLHLEGMMKVPNNKR 1485
QY 1485 GQGWDRKVIYLGSKVLIYDNEAREAGORPVEEPELCLPDGDVSIHGAVGASELANTAK 1544
Db 1486 GQGWDRKVIYLGSKVLIYDNEAREAGORPVEEPELCLPDGDVSIHGAVGASELANTAK 1545
QY 1545 A-----EKABADAKL 1554
Db 1546 ADVPYIILKMHESHPTTCPCRTLYLLA P S F PDKQRWVTAL E S V W A G R V S R E K A E A D A K L 1605
QY 1555 LGNSLLKLGDDRLDNKCTLPSDQVVLVGTBEGLYALNVKNLSLTHVPGIGAVQIYII 1614
Db 1606 LGNSLLKLGDDRLDNKCTLPSDQVVLVGTBEGLYALNVKNLSLTHVPGIGAVQIYII 1665
QY 1615 KDLKLLMTAGEBRALCLVDVKKVQSLAQSHLPAQPDISPNI FPAVKGCHLFAGAKTEN 1674
Db 1666 KDLKLLMTAGEBRALCLVDVKKVQSLAQSHLPAQPDISPNI FPAVKGCHLFAGAKTEN 1725
QY 1675 GLICICAMPKSVILRYNENLSKYCIKRIETSEPCSCIHTFNYSILIGTNKFIYIDMKQ 1734
Db 1726 GLICICAMPKSVILRYNENLSKYCIKRIETSEPCSCIHTFNYSILIGTNKFIYIDMKQ 1765
QY 1735 YTLLEFLDKNDHSLAPAVFAAASNSPVSIVQVNSAGQREYLLCFHFEGFVDSYGRS 1794
Db 1786 YTLLEFLDKNDHSLAPAVFAAASNSPVSIVQVNSAGQREYLLCFHFEGFVDSYGRS 1845
QY 1795 RTDDLKWSRLPLAFAYREPFLVFTHFNSLVIEIQARSSAGTTPARAYLDIPNRYLGPAL 1854

Db 1846 RTDDLKWSRLPLAFAYREPFLVFTHFNSLVIEIQARSSAGTTPARAYLDIPNRYLGPAL 1905
QY 1855 SSGAIYLAASSYQDKLRLVICCKGNLVKESGTEHHRGPTSSSSPNKRGPPTYNEHITKRVA 1914
Db 1906 SSGAIYLAASSYQDKLRLVICCKGNLVKESGTEHHRGPTSSSSPNKRGPPTYNEHITKRVA 1965
QY 1915 SSPAPPGPSHPREPSTPHRYREGRTELRDKSPGRDLEREKSPGMLSTRERSRGLF 1974
Db 1966 SSPAPPGPSHPREPSTPHRYREGRTELRDKSPGRDLEREKSPGMLSTRERSRGLF 2025
QY 1975 EDSRGRGLPAGAVRTPLSQVKNKGRGQSA 2002
Db 2026 EDSRGRGLPAGAVRTPLSQVKNKWDQSS 2053

RESULT 5

US-09-964-956-9
; Sequence 9, Application US/09964956
; Publication No. US20040043926A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964,956
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276,667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294,823
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304,868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2066

; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-964-956-9									
Query Match 95.4%; Score 10007; DB 12; Length 2066;									
Best Local Similarity 95.6%; Pred. No. 0;									
Matches 1979; Conservative 3; Mismatches 7; Indels 82; Gaps 4;									
QY	1	MLKFKYGARNPLDAGAAEPIASRASRLNFFGCKPFFMTQQQMSPLSREGILDALFVLFE	60						
DB	1	MLKFKYGARNPLDAGAAEPIASRASRLNFFGCKPFFMTQQQMSPLSREGILDALFVLFE	60						
QY	61	ECSQPALMKI KHVSFNVRKYSPTTIAELOLPQSAKDFFVRSIVGCGHFAEYVVREKATG	120						
DB	61	ECSQPALMKI KHVSFNVRKYSPTTIAELOLPQSAKDFFVRSIVGCGHFAEYVVREKATG	120						
QY	121	DIYAMKWKKALLAQEQVSPFEEERNILSRSTSPWIPOLQYAFODKNHLVLMBEYQPGG	180						
DB	121	DIYAMKWKKALLAQEQVSPFEEERNILSRSTSPWIPOLQYAFODKNHLVLMBEYQPGG	180						
QY	181	DLLSLINRYEDQIDENLITQFYLAELILAVHSVHLNGYVHRDIPKENILVDRTHIKLVDF	240						
DB	181	DLLSLINRYEDQIDENLITQFYLAELILAVHSVHLNGYVHRDIPKENILVDRTHIKLVDF	240						
QY	241	GSAAKWNSKMNNAKLPIGTPTDYMAPEVLTVMNGDGKGTGGLDCDWWSGVVIAYEMIYGR	300						
DB	241	GSAAKWNSK - VNAKLPIGTPTDYMAPEVLTVMNGDGKGTGGLDCDWWSGVVIAYEMIYGR	299						
QY	301	SPFAGTSGARTNNIMNFORFLKFPDDPKVSSDFLDLIQSLLCGQKBLKXFEGLCCHPPF	360						
DB	300	SPFAGTSGARTNNIMNFORFLKFPDDPKVSSDFLDLIQSLLCGQKBLKXFEGLCCHPPF	359						
QY	361	SKIDWNNIRNSPPFPVPTLKSDDTSNFDPEPKNSWSSPCQLSPSGFSGEEPLFVGFS	420						
DB	360	SKIDWNNIRNAPPPFPVPTLKSDDTSNFDPEPKNSWSSPCQLSPSGFSGEEPLFVGFS	419						
QY	421	YSKALGILGRSSVSGLDSPAKTSMKELLIKSKELQSDQKCHKMEQEMTRLHRVS	480						
DB	420	YSKALGILGRSSVSGLDSPAKTSMKELLIKSKELQSDQKCHKMEQEMTRLHRVS	479						
QY	481	EVEAVLSQKEVELKASEQRSLEODLATYITECSSLRSLRLEOARMVYSQEDDKALQLLH	540						
DB	480	EVEAVLSQKEVELKASEQRSLEODLATYITECSSLRSLRLEOARMVYSQEDDKALQLLH	539						
QY	541	DIREQSRKLQEIKEQYQAOQVEEMLMNQLEEDLVARRSDLYESELRESRLAAEFK	600						
DB	540	DIREQSRKLQEIKEQYQAOQVEEMLMNQLEEDLVARRSDLYESELRESRLAAEFK	599						
QY	601	RKATECOHKLKAKOQGPVEGEYAKLEKINAEQOLKIQELQEKLEAKVKAATEATELLQ	648						
DB	600	RKATECOHKLKAKOQGPVEGEYAKLEKINAEQOLKIQELQEKLEAKVKAATEATELLQ	659						
QY	649	---AKERAERELEKLOQRNEDSSEGIKKLVFAEBERRHSLNKKVRLTMMERENRLKDD	704						
DB	660	NTRQAKERAERELEKLOQRNEDSSEGIKKLVFAEBERRHSLNKKVRLTMMERENRLKDD	719						
QY	705	IQTQSQIQOQWADKILEEKEHREAQVSAQHLVHLKQKEQHYEKKIKVLDNQIKKDLAD	764						
DB	720	IQTQSQIQOQWADKILEEKEHREAQVSAQHLVHLKQKEQHYEKKIKVLDNQIKKDLAD	779						
QY	765	KETLENMQRHEEAEHKGKILSEOKAMINANDSKIRSLRQIVELSEANKLAANSLSFT	824						
DB	780	KETLENMQRHEEAEHKGKILSEOKAMINANDSKIRSLRQIVELSEANKLAANSLSFT	839						
QY	825	QRNMKAQEMISLRLQKQKPYLETQAGKLEAQNKRKLEEQLEKISHQCHSDKNRLLELETRL	884						
DB	840	QRNMKAQEMISLRLQKQKPYLETQAGKLEAQNKRKLEEQLEKISHQCHSDKNRLLELETRL	899						
QY	885	REVSLHEBEQKLEIKRQLTQLSLQEBRESQITALQAARAALESQLOAKTELBETTHAEA	944						
DB	900	REVSLHEBEQKLEIKRQLTQLSLQEBRESQITALQAARAALESQLOAKTELBETTHAEA	959						
QY	945	EBEIQALTAHRDEIQKFDALRNSCTVITDLEEQNLQTNEDNAELNNQNFYLSQQLDEAS	1004						

DB	960	EBEIQALTAHRDEIQKFDALRNSCTVITDLEEQNLQTNEDNAELNNQNFYLSQQLDEAS	1019						
QY	1005	GANDEIVQLRSEVDHLREITEREMQLTSQKQWMEALKTTCTMLBEOVMDEALNDELLE	1064						
DB	1020	GANDEIVQLRSEVDHLREITEREMQLTSQKQWMEALKTTCTMLBEOVMDEALNDELLE	1079						
QY	1065	KEROWEAWRVLGDEKSGQFCRVREIQORMLDITKQSRARADQORITRSQVVEAVKEHKA	1124						
DB	1080	KEROWEAWRVLGDEKSGQFCRVREIQORMLDITKQSRARADQORITRSQVVEAVKEHKA	1139						
QY	1125	EIALAQALKEOKLKABSLSDKLNLDLEKKHAMEMARSQOKLETERELKORLLBEQAK	1184						
DB	1140	EIALAQALKEOKLKABSLSDKLNLDLEKKHAMEMARSQOKLETERELKORLLBEQAK	1199						
QY	1185	LQOQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYOLENIQVLYSHKVKVGEFTISQ	1244						
DB	1200	LQOQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYOLENIQVLYSHKVKVGEFTISQ	1259						
QY	1245	QTKLIDFLQAKMCPAKKKKGLFSRRKEDPALPTQVPLQVNEKLKALEKAKCAELEEA	1304						
DB	1260	QTKLIDFLQAKMCPAKKKK-----VPLQVNEKLKALEKAKCAELEEA	1304						
QY	1305	LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSS	1364						
DB	1305	LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSS	1364						
QY	1365	RRKESSTPEFSRRLKERMHNI PHRFNVGLNVRATKCAVCLDTVHFGROASKLEQCVW	1424						
DB	1365	RRKESSTPEFSRRLKERMHNI PHRFNVGLNVRATKCAVCLDTVHFGROASKLEQCVW	1424						
QY	1425	CHPKCSTCLPATCGLPAEYATHTFACRDKNMSPGLQTKPESSSLHLEGMKVPNNKR	1484						
DB	1425	CHPKCSTCLPATCGLPAEYATHTFACRDKNMSPGLQTKPESSSLHLEGMKVPNNKR	1484						
QY	1485	GOQGWDRKYTVLPGSKVLIYDNEARAGORPVEEFELCLPDGVSIHGAVGASELANTAK	1544						
DB	1485	GOQGWDRKYTVLPGSKVLIYDNEARAGORPVEEFELCLPDGVSIHGAVGASELANTAK	1544						
QY	1545	A-----EKAEADAKL	1554						
DB	1545	ADVPYILKMESHPTTCWPGORTLYLAPSPDKQWWTALLESVVAGGRVSRREKAEADAKL	1604						
QY	1555	LGNLSLLKLEGGDRDLMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQIYII	1614						
DB	1605	LGNLSLLKLEGGDRDLMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQIYII	1664						
QY	1615	KDLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKVGCHLFGAGKIEN	1674						
DB	1665	KDLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKVGCHLFGAGKIEN	1724						
QY	1675	GLCICAAMPKSVILRYNENLSKYCIKKEITETSEPCSCIHFTNYSILIGTNKFEYIDMKQ	1734						
DB	1725	GLCICAAMPKSVILRYNENLSKYCIKKEITETSEPCSCIHFTNYSILIGTNKFEYIDMKQ	1784						
QY	1735	YTLLEFIDKXNDHSLAPAVFAASSNSPVSIVQVNSAGQREYLLCPHFQVFDVSYGRRS	1794						
DB	1785	YTLLEFIDKXNDHSLAPAVFAASSNSPVSIVQVNSAGQREYLLCPHFQVFDVSYGRRS	1844						
QY	1795	RTDCLKWSRLPLAFAYREPVLFTVHNSLEVIIEIQAESASAGTPARAYLDIPNPRYLGPAL	1854						
DB	1845	RTDCLKWSRLPLAFAYREPVLFTVHNSLEVIIEIQAESASAGTPARAYLDIPNPRYLGPAL	1904						
QY	1855	SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITKRVA	1914						
DB	1905	SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITKRVA	1964						
QY	1915	SSAPPEGSPHPEPSTPHRYREGTRELPRDKSPGRLEREKSPGRLMSTPERSRSLF	1974						
DB	1965	SSAPPEGSPHPEPSTPHRYREGTRELPRDKSPGRLEREKSPGRLMSTPERSRSLF	2024						
QY	1975	EDSSRGRPLPAGAVRTPLSQVNKGRGQASQV	2005						

Db 2025 EDSRGRPLPAGAVRTPLSQVNVKVRQHSACV 2055

RESULT 6

US-10-262-511-14

; Sequence 14, Application US/10262511

; Publication No: US20040038223A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glenda

; APPLICANT: Millst, Isabelle

; APPLICANT: Peyman, John A.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Ju, Jingfang

; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Patturajan, Meera

; APPLICANT: Szytek, Kimberly A.

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Ort, Tatiana

; APPLICANT: Gorman, Linda

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Anderson, David W.

; APPLICANT: Zhong, Wei

; APPLICANT: Catterton, Elina

; APPLICANT: Ji, Weizhen

; APPLICANT: Miller, Charles E.

; APPLICANT: Rastelli, Luca

; APPLICANT: Stone, David J.

; APPLICANT: Pena, Carol E. A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Rothenberg, Mark E.

; APPLICANT: Léach, Martin D.

; APPLICANT: Adege, Michele L.

; APPLICANT: Berghs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-462C

; CURRENT APPLICATION NUMBER: US/10/262,511

; CURRENT FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/373,815

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,642

; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/373,260

; PRIOR FILING DATE: 2002-04-17

; PRIOR APPLICATION NUMBER: 60/373,826

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,435

; PRIOR FILING DATE: 2001-10-05

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 439

; SOFTWARE: CuraSeqList version 0.1

; SEQ ID NO 14

; LENGTH: 2066

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-262-511-14

Query Match 95.4%; Score 10007; DB 12; Length 2066;

Best Local Similarity 95.6%; Pred. No. 0;

Matches 1979; Conservative 3; Mismatches 7; Indels 82; Gaps 4;

QY	1	MLKFKYGARPNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALVLFPE	60
DB	1	MLKFKYGARPNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALVLFPE	60
QY	61	ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKAFVRSVLVCCGHFAEVQVVRKATG	120
DB	61	ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKAFVRSVLVCCGHFAEVQVVRKATG	120
QY	121	DIYAMKVMKKALLAQBOVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGG	180
DB	121	DIYAMKVMKKALLAQBOVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLVLMEEYQPGG	180
QY	181	DLISLNRVEDQDENLIQFYLAELILAVHSVHLMGVHRDIKPNILVDRTGHIKLVDF	240
DB	181	DLISLNRVEDQDENLIQFYLAELILAVHSVHLMGVHRDIKPNILVDRTGHIKLVDF	240
QY	241	GSAAKMNSKVMNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCCDWWSVGVIAEVIYGR	300
DB	241	GSAAKMNSK-VNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCCDWWSVGVIAEVIYGR	299
QY	301	SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFDLIQSLLCQKERLKFEGLCCHPFF	360
DB	300	SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFDLIQSLLCQKERLKFEGLCCHPFF	359
QY	361	SKIDWNNIRNSPPFPVPTLKSDDDTSNFDEPKNSWSSPCOLSPSGSGEELPFVGS	420
DB	360	SKIDWNNIRNAPPPFPVPTLKSDDDTSNFDEPKNSWSSPCOLSPSGSGEELPFVGS	419
QY	421	YSKALGILGRSESVSGLDSPAKTSSMEKKLLIKSKELQSDQDKCHKMEQMTLHRRYS	480
DB	420	YSKALGILGRSESVSGLDSPAKTSSMEKKLLIKSKELQSDQDKCHKMEQMTLHRRYS	479
QY	481	EVBAVLQKEVELKASRTQSRSLLEODLATVITCSLSIKSLEQARMEVSGEDDKALQLLH	540
DB	480	EVBAVLQKEVELKASRTQSRSLLEODLATVITCSLSIKSLEQARMEVSGEDDKALQLLH	539
QY	541	DIREQSKLOEIKQEYQAOVEEMRLMNOLEEDLVSAARRSDLYSELSRESLAABEFK	600
DB	540	DIREQSKLOEIKQEYQAOVEEMRLMNOLEEDLVSAARRSDLYSELSRESLAABEFK	599
QY	601	RKATECOHKLKADQKQPEVGYAKLEKINAFQOLKIQELQKLEKAKVASTATELLQ	648
DB	600	RKATECOHKLKADQKQPEVGYAKLEKINAFQOLKIQELQKLEKAKVASTATELLQ	659
QY	649	-----AKERAERELEKLNREDSSEGIRKLVAEERRHS-ENKVKLETWRENERLKKD	704
DB	660	NIRQAKERARELEKLNREDSSEGIRKLVAEERRHS-ENKVKLETWRENERLKKD	719
QY	705	IQTKSQIQQMADKIILEEEKHREAQVSAQHLVHLKQKEQHYEEKIKVLNQIKKDLAD	764
DB	720	IQTKSQIQQMADKIILEEEKHREAQVSAQHLVHLKQKEQHYEEKIKVLNQIKKDLAD	779
QY	765	KETLENNMQHREBEAHEKGIILSEQAMINAMDSKIRSLRQIVELSEANKLAANSFLT	824
DB	780	KETLENNMQHREBEAHEKGIILSEQAMINAMDSKIRSLRQIVELSEANKLAANSFLT	839
QY	825	QRNWKAEEMISELROOKFYLETQAGKLEAQRNKLBEQLEKISHQDHSQDKNRLLETRL	884
DB	840	QRNWKAEEMISELROOKFYLETQAGKLEAQRNKLBEQLEKISHQDHSQDKNRLLETRL	899
QY	885	REVSLEHEEOKLEIKRQLTBLQSLQERESQALQAAARALQSOLQAKTELESETTAA	944
DB	900	REVSLEHEEOKLEIKRQLTBLQSLQERESQALQAAARALQSOLQAKTELESETTAA	959
QY	945	EEBIQALTARDBIQRFDAIRNSCTVITDLERQNLQLTEDNAELNNQNFYLSKQLEAS	1004
DB	960	EEBIQALTARDBIQRFDAIRNSCTVITDLERQNLQLTEDNAELNNQNFYLSKQLEAS	1019
QY	1005	GANDEIVQLRSEVDHLRREITEREMQLTQSQQTMWALKTTCTMLBEQVMDLEALNDELLE	1064
DB	1020	GANDEIVQLRSEVDHLRREITEREMQLTQSQQTMWALKTTCTMLBEQVMDLEALNDELLE	1079

1065 KERQWAWRSVLGDEKSPQCHVRELQRLMDTEKQSRAPADORIITESTQVVELAVKEHA 1124
1080 KERQWAWRSVLGDEKSPQCHVRELQRLMDTEKQSRAPADORIITESTQVVELAVKEHA 1139
1125 EIALAQALKEQKLKAEISLNDLEKHALENNARSLOOKLETERELKORLLEBOAK 1184
1140 EIALAQALKEQKLKAEISLNDLEKHALENNARSLOOKLETERELKORLLEBOAK 1199
1185 LQQQMDLQKHIFRUTQGLQAEALDRADLLKTERSDLEYOLENIQVLYSHEKYMEGTISQ 1244
1200 LQQQMDLQKHIFRUTQGLQAEALDRADLLKTERSDLEYOLENIQVLYSHEKYMEGTISQ 1259
1245 QTKLIDFLQAKWDQPAKKKGLFSRKEDPALPTQVPLQYNELKLALEKEKARCALDEEA 1304
1260 QTKLIDFLQAKWDQPAKKK-----VPLQYNELKLALEKEKARCALDEEA 1304
1305 LQKTRIELRSAREEAHRKATDHPHPSTPATARQQAIAKSAIVRSEPHQPSAMSLAPRSS 1364
1305 LQKTRIELRSAREEAHRKATDHPHPSTPATARQQAIAKSAIVRSEPHQPSAMSLAPRSS 1364
1365 RKESSTPEFRRRLKERHNNIPIHRFNVLGNWRAKCAVCLDTVHFGROASKCLECOVM 1424
1365 RKESSTPEFRRRLKERHNNIPIHRFNVLGNWRAKCAVCLDTVHFGROASKCLECOVM 1424
1425 CHPKSTCULPATCGUPAEVATHFTFAFCRDKVNSPQLOTKEPSSILHLEGMKVPNNKR 1484
1425 CHPKSTCULPATCGUPAEVATHFTFAFCRDKVNSPQLOTKEPSSILHLEGMKVPNNKR 1484
1485 GQGGWDRKVIIVLEGSKVLIIYDNEAREAGQRPVEEFELCLPDGDSVHGAAGASELANAK 1544
1485 GQGGWDRKVIIVLEGSKVLIIYDNEAREAGQRPVEEFELCLPDGDSVHGAAGASELANAK 1544
1545 A-----EKAADAKL 1554
1545 ADVPIYLMKESHPHTTCWPGRTLYLLAPSPDKQWVTALESVAGRVSRKEAADAKL 1604
1555 LGNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLNLSLTHVPGIGAVFQIYII 1614
1605 LGNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLNLSLTHVPGIGAVFQIYII 1664
1615 KOLEKLMTAGBERALCLVDVKKVQSLAOSHLPAPQDISPNIPEAVKGCHLFGAGKIEN 1674
1665 KOLEKLMTAGBERALCLVDVKKVQSLAOSHLPAPQDISPNIPEAVKGCHLFGAGKIEN 1724
1675 GLCICAAMPKVVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSGILGTNKFYEIDMKQ 1734
1725 GLCICAAMPKVVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSGILGTNKFYEIDMKQ 1784
1735 YTLBFLDKNDHSLAPAVPAASNSFPVSVIVQNSAGQREYLLCFHFEGFVVDVSYGRRS 1794
1785 YTLBFLDKNDHSLAPAVPAASNSFPVSVIVQNSAGQREYLLCFHFEGFVVDVSYGRRS 1844
1795 RTDDLKWSRLPLAFVAREPVLVTHFNSLEVIQIARSSAGTPARAYLDIPNRYLGPAL 1854
1845 RTDDLKWSRLPLAFVAREPVLVTHFNSLEVIQIARSSAGTPARAYLDIPNRYLGPAL 1904
1855 SSGAIYASSYQDKLRVICCKGNLVKESGTEHHRGPFSTRSSPNKRGPPTTYNEHITKRAVA 1914
1905 SSGAIYASSYQDKLRVICCKGNLVKESGTEHHRGPFSTRSSPNKRGPPTTYNEHITKRAVA 1964
1915 SFPAPPEGSHPRESTPHRYREGTELRDSDRSPRPLEREXSPGMLSTRERSPGRLF 1974
1965 SFPAPPEGSHPRESTPHRYREGTELRDSDRSPRPLEREXSPGMLSTRERSPGRLF 2024
1975 EDSSRGRPLPAGAVRTPLSQVKNKGRGQASQV 2005
2025 EDSSRGRPLPAGAVRTPLSQVKNKGRHSEACV 2055

RESULT 7
US-09-964-956-11
; Sequence 11, Application US/09964956
; Publication No. US20040043926A1

GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964,956
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276,667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294,823
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304,868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-964-956-11

Query Match 95.4%; Score 10005; DB 12; Length 2053;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;

QY 1 MLKPKYGARNPLDGAABPIASRASRLNLFQCKPPPTMQQMSPLSRREGILDALFVLFE 60
DB 1 MLKPKYGARNPLDGAABPIASRASRLNLFQCKPPPTMQQMSPLSRREGILDALFVLFE 60
QY 61 ECSOPALMKIKHVSNFVKYSYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120
DB 61 ECSOPALMKIKHVSNFVKYSYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120
QY 121 DIYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQYAFQDKKHHLYLMEEYOPGG 180
DB 121 DIYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQYAFQDKKHHLYLMEYOPGG 180

181 DLSLLNRYEDQDLENLQFYLAELILAVSHVLMGYVHRDIKPENILVDRTHIKLVD 240
181 DLSLLNRYEDQDLENLQFYLAELILAVSHVLMGYVHRDIKPENILVDRTHIKLVD 240
241 GSAKKNNSKNNVAKLPIGTPDYMAEVLTVMGDQKGTIYGLDCDWWSGVYAYEMIYGR 300
241 GSAKKNNSK-NVAKLPIGTPDYMAEVLTVMGDQKGTIYGLDCDWWSGVYAYEMIYGR 299
301 SPAEGTSARTFNIMNFQFLFPDDPKVSSDFLDLIQSLCCQKRLKFEGLCCHPFF 360
300 SPAEGTSARTFNIMNFQFLFPDDPKVSSDFLDLIQSLCCQKRLKFEGLCCHPFF 359
361 SKIDWNIRNSPPFPVPTLTKSDDDTNSFNDEPEKNSWSSPCOLSPSGFSGBELPVGFS 420
360 SKIDWNIRNAPPFPVPTLTKSDDDTNSFNDEPEKNSWSSPCOLSPSGFSGBELPVGFS 419
421 YSVALGILGRSESVVSGLDSPAKTSMEKKLILKSKELQSDQKCHKMQEMLRHRVS 480
420 YSVALGILGRSESVVSGLDSPAKTSMEKKLILKSKELQSDQKCHKMQEMLRHRVS 479
481 EVBAVLSQKEVELKASQTSRSLLEQDLATVITECSSLSKRSLEQARMEVSDQDDKALQLH 540
480 EVBAVLSQKEVELKASQTSRSLLEQDLATVITECSSLSKRSLEQARMEVSDQDDKALQLH 539
541 DIREQSKLOEIKBOEYQAOVEWRLMNQLEBDLYSARRSDLYESELRESRLAAEFK 600
540 DIREQSKLOEIKBOEYQAOVEWRLMNQLEBDLYSARRSDLYESELRESRLAAEFK 599
601 RKATECOHLLKAKDQKPVGVYAKLEKINAEQQLKIQELQEKLEKA----- 648
600 RKATECOHLLKAKDQKPVGVYAKLEKINAEQQLKIQELQEKLEKAKVKASTEATLQ 659
649 -----AKERABRELEKUNRDSSEGTRKKLVEABERRHSLNKVKLETMERENRLKDD 704
660 NIRAQERABRELEKUNRDSSEGTRKKLVEABERRHSLNKVKLETMERENRLKDD 719
705 IQTKSQOIQOMADKILLEEKHREAOVSQHLVEHLKQEHVEEKIKVLDNQIKKDLAD 764
720 IQTKSQOIQOMADKILLEEKHREAOVSQHLVEHLKQEHVEEKIKVLDNQIKKDLAD 779
765 KETLENMQRHEBEAHEKGIUSEQKAMINAMDSKIRSLQRIVELSEANKLAANSUFT 824
780 KETLENMQRHEBEAHEKGIUSEQKAMINAMDSKIRSLQRIVELSEANKLAANSUFT 839
825 ORNKQAQEMISELROOKFVLETOAQKLEAQNKLSEOLEKISHOODHSDKNRLLLETRL 884
840 ORNKQAQEMISELROOKFVLETOAQKLEAQNKLSEOLEKISHOODHSDKNRLLLETRL 899
885 REVSLEHEEQKLEKRLQTELQLSQERESQLTALQARAALLESQLRQAKTELETTAA 944
900 REVSLEHEEQKLEKRLQTELQLSQERESQLTALQARAALLESQLRQAKTELETTAA 959
945 BEBIOALTARHDEIQKFPALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSKOLDEAS 1004
960 BEBIOALTARHDEIQKFPALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSKOLDEAS 1019
1005 GANDEIVQLRSEVDHLRREITEREMOLTSOKOTMEALKTTCTMLBEOQVMDLEALDELLE 1064
1020 GANDEIVQLRSEVDHLRREITEREMOLTSOKOTMEALKTTCTMLBEOQVMDLEALDELLE 1079
1065 KERQWEANRSLVGDSEKQFECRVREIQRLMDTEKQSRARADQIRITESROVVELAVKEHA 1124
1080 KERQWEANRSLVGDSEKQFECRVREIQRLMDTEKQSRARADQIRITESROVVELAVKEHA 1139
1125 EILALQOALKKEOKLKAESLSDKNDLEKKHMLENNARSLQOKLETERELKORLLEEQAK 1184
1140 EILALQOALKKEOKLKAESLSDKNDLEKKHMLENNARSLQOKLETERELKORLLEEQAK 1199
1185 LQOQMDLQKNHIFRLTOGQOALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1244
1200 LQOQMDLQKNHIFRLTOGQOALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1259
1245 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVPLQYNELKLALEKEKARCAELEEA 1304

1260 QTKLIDFLQAKMDQPAKKKK-----VPLQYNELKLALEKEKARCAELEEA 1304
1305 LOKTRIELRSAREBAHRKATDHPSTPATARQOIAMSAIVSRPEHQPSAMSLAPSS 1364
1305 LOKTRIELRSAREBAHRKATDHPSTPATARQOIAMSAIVSRPEHQPSAMSLAPSS 1364
1365 RRKESSTPEEFSSRLKERMHEHNPBRENVGLNMRATKCAVCLDTVHFGRQASKLEQVM 1424
1365 RRKESSTPEEFSSRLKERMHEHNPBRENVGLNMRATKCAVCLDTVHFGRQASKLEQVM 1424
1425 CHPKCSTCLPATCGLPAYATHTEAFRCRDMNSPGLCTKEPSSSLHLEGMKVPNNKR 1484
1425 CHPKCSTCLPATCGLPAYATHTEAFRCRDMNSPGLCTKEPSSSLHLEGMKVPNNKR 1484
1485 GOQWDRKXYIVLEGSKVLIYDNEAREAGQRPVEFEFCLCLPDGDSVHGAAGASELANTAK 1544
1485 GOQWDRKXYIVLEGSKVLIYDNEAREAGQRPVEFEFCLCLPDGDSVHGAAGASELANTAK 1544
1545 A-----EKAEDAKL 1554
1545 ADVPYILKMHSHPTTCMPGRTLYLLAPSPDKQWVTALESVVAGGRVSRKAEADAKL 1604
1555 LGNSLLKLEGGDRDDNMCTLPFSQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1614
1605 LGNSLLKLEGGDRDDNMCTLPFSQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1664
1615 KXLEKLLMIAGEERALCLVDVKKVQSLAQSHLPAQDISPNIFEAVKGCFLFGAGKIEN 1674
1665 KXLEKLLMIAGEERALCLVDVKKVQSLAQSHLPAQDISPNIFEAVKGCFLFGAGKIEN 1724
1675 GLCICAAMPKSVTLRNLNENLKYCIKREIETSEPCSHFTNYSLIGTNKFEYEDMKQ 1734
1725 GLCICAAMPKSVTLRNLNENLKYCIKREIETSEPCSHFTNYSLIGTNKFEYEDMKQ 1784
1735 YTLFEEDLKNDSHSLAPAVFAASNSFPVSIQVNSAGOREEYLLCFHFGVFDVSYGRS 1794
1785 YTLFEEDLKNDSHSLAPAVFAASNSFPVSIQVNSAGOREEYLLCFHFGVFDVSYGRS 1844
1795 RTDDLKWSRLPLAFAYREPVLVTHFNSLEVIIEIQAASSAGTPARAYLDIPNRYLGPAL 1854
1845 RTDDLKWSRLPLAFAYREPVLVTHFNSLEVIIEIQAASSAGTPARAYLDIPNRYLGPAL 1904
1855 SSGAIYLASSYQDKRLVICCKGNLKVESGTEHHRGPGSTSRSSPNKRGPTTYNEHITKRV 1914
1905 SSGAIYLASSYQDKRLVICCKGNLKVESGTEHHRGPGSTSRSSPNKRGPTTYNEHITKRV 1964
1915 SSPAPPGSPHREPPSTPHRYREGTRELDRDKSGFGRPLEREKSPGRMLSTRERSRSGRLF 1974
1965 SSPAPPGSPHREPPSTPHRYREGTRELDRDKSGFGRPLEREKSPGRMLSTRERSRSGRLF 2024
1975 EDSRGRPLPAGAVRTPLSQVKNKRGQSA 2002
2025 EDSRGRPLPAGAVRTPLSQVKNKWDQSS 2052

RESULT 8

US-10-262-511-2
; Sequence 2, Application US/10262511
; Publication No. US20040039223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.


```

; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-2

Query Match          95.4%; Score 10005; DB 12; Length 2053;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;

Qy 1 MLKFKYGARNPLDAGAAPRIASRSLNLFQGGKPPFTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFKYGARNPLDAGAAPRIASRSLNLFQGGKPPFTQQQMSPLSREGILDALFVLFE 60

Qy 61 ECSQPALMKIKHVSFNPKYSOTIAELOQPSAKDFEVRSLVGCCHFAEVQVREKATG 120
Db 61 ECSQPALMKIKHVSFNPKYSOTIAELOQPSAKDFEVRSLVGCCHFAEVQVREKATG 120

Qy 121 DIYAKVMKKKALLAQEQVSFEERNILSRSTSWIPOLOVAFODKKNHLYLMEYQPGG 180
Db 121 DIYAKVMKKKALLAQEQVSFEERNILSRSTSWIPOLOVAFODKKNHLYLMEYQPGG 180

Qy 181 DLLSLNRYEDQDENLIQFYLAELILAVHSLMGYVHRDIKPNILVDRGTGHIKLVD 240
Db 181 DLLSLNRYEDQDENLIQFYLAELILAVHSLMGYVHRDIKPNILVDRGTGHIKLVD 240

Qy 241 GSAAKMNSNMVNAKLPIGTPTDYNAPEVLTVNMGDKGTGYLDCDWSVGVIAYEMIYGR 300
Db 241 GSAAKMNSNK-VNAKLPIGTPTDYNAPEVLTVNMGDKGTGYLDCDWSVGVIAYEMIYGR 299

```

```

Qy 301 SPRAEGTSAETNNIMNFQRLKFPDDPKVSSDFLDLIQSLLCGQKRLKFEGLCCHPEF 360
Db 300 SPFAEGTSARTFNIMNFQRLKFPDDPKVSSDFLDLIQSLLCGQKRLKFEGLCCHPEF 359

Qy 361 SKIDMNNIRNSPPFPVPTLKSDDDTSNFDEPKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Db 360 SKIDMNNIRNAPFPVPTLKSDDDTSNFDEPKNSWVSSSPCQLSPSGFSGEELPFVGF 419

Qy 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKLEQSDQDKCHQWQEMTFLHRVS 480
Db 420 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKLEQSDQDKCHQWQEMTFLHRVS 479

Qy 481 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSILKRSLEQARMEVSDDKALQLLH 540
Db 480 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSILKRSLEQARMEVSDDKALQLLH 539

Qy 541 DIRQSKLOEIKQEQYQAOVEEMRLMMNOLBEDLVSAARRSDLYSELRESLIAAEFK 600
Db 540 DIRQSKLOEIKQEQYQAOVEEMRLMMNOLBEDLVSAARRSDLYSELRESLIAAEFK 599

Qy 601 RKATECOHKLKAKDQKPEVGYAKLEKINAEQQLKIQLQEKLEKA----- 648
Db 600 RKATECOHKLKAKDQKPEVGYAKLEKINAEQQLKIQLQEKLEKAVKASTEATELLQ 659

Qy 649 ----AKERAERELEKLQNRDSSSEGIKKLVEAEERRHSLKVKLETTWERENRLKDD 704
Db 660 NIRQAKERAERELEKLQNRDSSSEGIKKLVEAEERRHSLKVKLETTWERENRLKDD 719

Qy 705 IOTKSQOIQQMADKILEBEKHEBAQVSAOHLEVHLKQEOHYEELIKVLNDQIKKDLAD 764
Db 720 IOTKSQOIQQMADKILEBEKHEBAQVSAOHLEVHLKQEOHYEELIKVLNDQIKKDLAD 779

Qy 765 KETLENMVMQREHEBAHEKGIKILSEQKAMINAMSKIRSLRQIRIVELSEANKLAANSFLT 824
Db 780 KETLENMVMQREHEBAHEKGIKILSEQKAMINAMSKIRSLRQIRIVELSEANKLAANSFLT 839

Qy 825 QRNKAQEMISELRQOKFYLETQAGKLEAQNRLKEBQLEKISHQDSDKNRLLLETRL 884
Db 840 QRNKAQEMISELRQOKFYLETQAGKLEAQNRLKEBQLEKISHQDSDKNRLLLETRL 899

Qy 885 REVSLEHEBQKLELKROLTELQLSQERESQTLQAAARAALSQRLCAKTELEETFAEA 944
Db 900 REVSLEHEBQKLELKROLTELQLSQERESQTLQAAARAALSQRLCAKTELEETFAEA 959

Qy 945 EEEIQALTARDEIQRFKDALRNSCTVITDLEBQNLQEDNAELNNQNFYLSKQLDEAS 1004
Db 960 EEEIQALTARDEIQRFKDALRNSCTVITDLEBQNLQEDNAELNNQNFYLSKQLDEAS 1019

Qy 1005 GANDEIVOLSEVDHLREITEREMQLTSCQOTWEALKTCTMLEBOVMDLEALNDELLE 1064
Db 1020 GANDEIVOLSEVDHLREITEREMQLTSCQOTWEALKTCTMLEBOVMDLEALNDELLE 1079

Qy 1065 KERQWEAWRSVLGDEKSQFECRVRELQRMLDTEKQSRARADQRTESQVVELAVKEHKA 1124
Db 1080 KERQWEAWRSVLGDEKSQFECRVRELQRMLDTEKQSRARADQRTESQVVELAVKEHKA 1139

Qy 1125 EILALQALKEQKLEKASLSKDLNDEKKGAMLENNARSLOQKLETERELKORLLEQAK 1184
Db 1140 EILALQALKEQKLEKASLSKDLNDEKKGAMLENNARSLOQKLETERELKORLLEQAK 1199

Qy 1185 LQQOMDLQKNHIFRLTQGLQALDRDLKLTESDLEYQLENTQVLYSHEKVMGEGTISQ 1244
Db 1200 LQQOMDLQKNHIFRLTQGLQALDRDLKLTESDLEYQLENTQVLYSHEKVMGEGTISQ 1259

Qy 1245 QTKLIDFLQMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKALEKFKARCAELEEA 1304
Db 1260 QTKLIDFLQMDQPAKKK-----VPLQYNELKALEKFKARCAELEEA 1304

Qy 1305 LQKTRIELRSAREEAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAVSILAPSS 1364
Db 1305 LQKTRIELRSAREEAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAVSILAPSS 1364

```

QY	1365	RRKESSTPEEFSSRLXERMHNTI	PHRFNVGLNMRATKCAVCLDTVHFGQASKCLEQVM	1424
Db	1365	RRKESSTPEEFSSRLXERMHNTI	PHRFNVGLNMRATKCAVCLDTVHFGQASKCLEQVM	1424
QY	1425	CHPKCSTCLPATCGLPAEYATHTEAF	CRDKMNSPGLQTKESPSSUHLGWMKVPNNKR	1484
Db	1425	CHPKCSTCLPATCGLPAEYATHTEAF	CRDKMNSPGLQTKESPSSUHLGWMKVPNNKR	1484
QY	1485	GOQGWKRYIVLBSGSKVLIYDNR	BARAGQPVPEFELCLPDGVS IHGAVGASELANTAK	1544
Db	1485	GOQGWKRYIVLBSGSKVLIYDNR	BARAGQPVPEFELCLPDGVS IHGAVGASELANTAK	1544
QY	1545	A-----	-----EKAADAKL	1554
Db	1545	ADVPYILKMESHPTTCWPGRTIYL	LAPFPDQKRWVTALESVAVAGGRVSRKAADAKL	1604
QY	1555	LGNLLKLEGGDRDLNCTLPFSDQV	VLVGTBEGLYALANVLSLTHVPGIGAVFQIYII	1614
Db	1605	LGNLLKLEGGDRDLNCTLPFSDQV	VLVGTBEGLYALANVLSLTHVPGIGAVFQIYII	1664
QY	1615	KDLEKLLMINGERRALCLVDVKV	KOSLAQSHLPAQPDISPNI FBAVKCHLFGAGKIBN	1674
Db	1665	KDLEKLLMINGERRALCLVDVKV	KOSLAQSHLPAQPDISPNI FBAVKCHLFGAGKIBN	1724
QY	1675	GLGICAMPSPKVVILRYNENLSK	YCIKKEIETSEPCSCIHFTNYSILIGTNKIFYIDMKQ	1734
Db	1725	GLGICAMPSPKVVILRYNENLSK	YCIKKEIETSEPCSCIHFTNYSILIGTNKIFYIDMKQ	1784
QY	1735	YTLBFLDKNDHSLAPAVFAASNS	PPVSIQVNSAGQREBYLLCPHEFGVFDVSGRRS	1794
Db	1785	YTLBFLDKNDHSLAPAVFAASNS	PPVSIQVNSAGQREBYLLCPHEFGVFDVSGRRS	1844
QY	1795	RTDDLKWSRLPLAFAYREPILFV	THNSLEVEIQARSAGTPARAYLDIPNPRVILGPAT	1854
Db	1845	RTDDLKWSRLPLAFAYREPILFV	THNSLEVEIQARSAGTPARAYLDIPNPRVILGPAT	1904
QY	1855	SSGAIYLAGSYQDLRVICCKGNL	VKESGTEHHRGPSSTRSPNKGPPTYNEHITKVA	1914
Db	1905	SSGAIYLAGSYQDLRVICCKGNL	VKESGTEHHRGPSSTRSPNKGPPTYNEHITKVA	1964
QY	1915	SSPAPPEGSHPREPSTPHRYEGR	TBLRDKSPGRPLEREKSPGMLSTRERSPGRLF	1974
Db	1965	SSPAPPEGSHPREPSTPHRYEGR	TBLRDKSPGRPLEREKSPGMLSTRERSPGRLF	2024
QY	1975	EDSSRGRLPAGAVRTPLSQVNGR	QGSA 2002	
Db	2025	EDSSRGRLPAGAVRTPLSQVNGR	QSS 2052	

RESULT 9
US-10-017-216-4
; Sequence 4, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPILLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prob
; FILE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-4

Query Match 92.0%; Score 9656; DB 13; Length 2055;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;

QY	1	MLKFKYGARPLDAGAAEPIASRAS	LNLFPOGKPPFMTQQQSPISREGILDALFVIFE	60
Db	1	MLKFKYGRNPPPEASASEPIASRA	SNLNFPOGKPPPLMTQQQSAISREGMLDALFALFE	60
QY	61	BCSPALMKIGHNPNFYKYSDTIAE	LOQELQPSAKDFEVRSLVGCCHFAFVQVVRKATG	120
Db	61	BCSPALMKWGHVSSFKYSDTIAEL	RELOPSARDFEVRSLVGCCHFAFVQVVRKATG	120
QY	121	DIYAMKWKKALLAQQOVFFBEER	NILSRSTSPWIPQLOQYAFQDKNHHLYMBEETOPGG	180
Db	121	DYAMKMKKALLAQQOVFFBEER	NILSRSTSPWIPQLOQYAFQDKNHHLYMBEETOPGG	180
QY	181	DLLSILNRYEDQDENLIQFYLAEL	ILAVHSHVLMGVVHRDIKPENILVDRTGHILKLVDF	240
Db	181	DFLSILNRYEDQDENLIQFYLAEL	ILAVHSHVLMGVVHRDIKPENILVDRTGHILKLVDF	240
QY	241	GSAAKMNSKNWNAKLPIGTDYMA	PEVLTVMAGDGKTYGLDCDMMWSGVGIAYEMIYGR	300
Db	241	GSAAKMNSK-VDAKLPIGTDYMA	PEVLTVMAGDGKTYGLDCDMMWSGVGIAYEMIYGR	299
QY	301	SPFACETSARTFNINMFORFLK	FPDDPKYSSDFLDLIQSLICGOKERLKFGLCHPFF	360
Db	300	TPFTEGTSARTFNINMFORFLK	FPDDPKYSSBLLDLQLLQVQERLKFGLCHPFF	359
QY	361	SKIDWNIRNSPPFPVPTLAKSD	DDTNSNDFEPEKNSWVSSSPCQLSPSGFSGEELPFVGF	420
Db	360	ARTDWNIRNSPPFPVPTLAKSD	DDTNSNDFEPEKNSWAFILCVPAEPLAFSGEELPFVGF	419
QY	421	YSKALGILGRSES VVSGLDSP	AKTSMEKKLLIKSELQDSQDKCHMQEOMTFLHRRVS	480
Db	420	YSKALGILGRSES VVSGLDSP	AKTSMEKKLLIKSELQDSQDKCHMQEOMTFLHRRVS	479
QY	481	EVEAVLSQKEVELKASTQSL	LEQDLATYITBCSSLRKSLQARMEVSEDDKALQLLH	540
Db	480	EVEAVLSQKEVELKASTQSL	LEQDLATYITBCSSLRKSLQARMEVSEDDKALQLLH	539
QY	541	DIREQSKLQEIKEQEYQAOVEM	RMLMNQLEBDLSARRRDLSELESRESIAAEFFK	600
Db	540	DIREQSKLQEIKEQEYQAOVEM	RMLMNQLEBDLSARRRDLSELESRESIAAEFFK	599
QY	601	RKATECOHLLKAKDQKPEVGA	KLEKINAEQOLKIQELQEKLEA-----	648
Db	600	RKATECOHLLKAKDQKPEVGA	KLEKINAEQOLKIQELQEKLEKAVKASTENTELLQ	659
QY	649	---AKERAERELEKLNRED	SSGIRKKLVEAEERRHSLNKKVRLTMRERENLKDD	704
Db	660	NIRQAKERAERELEKLNRED	SSGIRKKLVEAEERRHSLNKKVRLTMRERENLKDD	719
QY	705	IOTKSOQIQOMADKILELEK	HRFAQVSAOHLVHLKQKECHVEEKKVLDNOIKKDLAD	764
Db	720	IOTKSEQIQOMADKILELEK	HRFAQVSAOHLVHLKQKECHVEEKKVLDNOIKKDLAD	779
QY	765	KETLENMQRHEBEAHEKGI	SEKAMINAMDSKIRSEQRIVELSEANKLAANSSFLT	824
Db	780	KESLENMQRHEBEAHEKGI	SEKAMINAMDSKIRSEQRIVELSEANKLAANSSFLT	839
QY	825	QRNKAQEEIMSELROOKFYLE	TOAGKLEAQNKLSEOLEKISHQDHSKVRILLETRL	884
Db	840	QRNKAQEEIMSELROOKFYLE	TOAGKLEAQNKLSEOLEKISHQDHSKVRILLETRL	899
QY	885	REVSLEHEEQKLELKROLTE	LOLSLOERESQTLAQARAALSESQLRQAOKTELEETAE	944
Db	900	REVSLEHEEQKLELKROLTE	LOLSLOERESQTLAQARAALSESQLRQAOKTELEETAE	959
QY	945	EEETQALTAHRDEIQRFDA	LRNSCTVITDLEBQLNQLTEDNAELNNQNFYLSKQJDEAS	1004
Db	960	EEETQALTAHRDEIQRFDA	LRNSCTVITDLEBQLNQLTEDNAELNNQNFYLSKQJDEAS	1019
QY	1005	GANDEIVQLRSEVDHLRRE	ITEREMOLTSQKQNEALKTTCTMLEBQVMDLEALNDELLE	1064
Db	1020	GANDEIVQLRSEVDHLRRE	ITEREMOLTSQKQNEALKTTCTMLEBQVMDLEALNDELLE	1079

```
QY 1065 KERQEWAMRSVLGDEKQSECEVRBLQRLMDTEKQSRARADORITQESQVVELAVKEHA 1124
Db 1080 KERQEWAMRSVLGDEKQSECEVRBLQRLMDTEKQSRARADORITQESQVVELAVKEHA 1139
QY 1125 EILALQOALKKEOKLKAESLSDKLNLEKXHAMLENWARSLOQKLETERELKORLEEOAK 1184
Db 1140 EILALQOALKKEOKLKAESLSDKLNLEKXHAMLENWARSLOQKLETERELKORLEEOAK 1199
QY 1185 LQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSLEYQLENIOVLYSHEKYMGEGTISQ 1244
Db 1200 LQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSLEYQLENIOVLYSHEKYMGEGTISQ 1259
QY 1245 QTKLIDFLOAKWDQPAKXKGLFSRKEDPALPTQVPOYNELKLALEKEKARCALIEEA 1304
Db 1260 QTKLIDFLOAKWDQPAKXKGLFSRKEDPALPTQVPOYNELKLALEKEKARCALIEEA 1304
QY 1305 LOKTRIELRSABEAAHRKATDHPHSTPATARQOIAMSIVRSPHOPFSAMSLAPSS 1364
Db 1305 LOKTRIELRSABEAAHRKATDHPHSTPATARQOIAMSIVRSPHOPFSAMSLAPSS 1364
QY 1365 RKESSTPEFGRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTHVFGROASKCLECOVM 1424
Db 1365 RKESSTPEFGRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTHVFGROASKCLECOVM 1424
QY 1425 CHPKSTCLPATCGLPAEVATHTFAFCRDKVNSPCLQTKPESSLSHLHLEGMKVPRNNKR 1484
Db 1425 CHPKSTCLPATCGLPAEVATHTFAFCRDKVNSPCLQTKPESSLSHLHLEGMKVPRNNKR 1484
QY 1485 GQGWDRKYIVLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544
Db 1485 GQGWDRKYIVLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544
QY 1545 A-----EKAADAKL 1554
Db 1545 ADVPYLKMSHPHTTCWPGRTLYLAPSPQKQWVTALESVAGRVSRKEADAKL 1604
QY 1555 LGNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFIYII 1614
Db 1605 LGNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFIYII 1664
QY 1615 KOLEKLLTAGBERALCLVDVKVKVQSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN 1674
Db 1665 KOLEKLLTAGBERALCLVDVKVKVQSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN 1724
QY 1675 GLCICAMPKSKVILRYNENISKYCIKREIETSEPCSCIHFTNYSILITGNKFYIDMKQ 1734
Db 1725 SUCICAMPKSKVILRYNENISKYCIKREIETSEPCSCIHFTNYSILITGNKFYIDMKQ 1784
QY 1735 YLLEEFLDKNDHSLAPAVFAASNSGFPVSIQVNSAGQREYLLCFHBEFGVFDVSYGRRS 1794
Db 1785 YLLEEFLDKNDHSLAPAVFAASNSGFPVSIQVNSAGQREYLLCFHBEFGVFDVSYGRRS 1844
QY 1795 RTDDLKWSRLPLAFAYREPLYFVTHNSLEVTEIQARSSAGTPAPAYLIDIPNRYLGPAI 1854
Db 1845 RTDDLKWSRLPLAFAYREPLYFVTHNSLEVTEIQARSSAGTPAPAYLIDIPNRYLGPAI 1904
QY 1855 SSGAIYLAASYQDKLKVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPTTYNEHITKVA 1914
Db 1905 SSGAIYLAASYQDKLKVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPTTYNEHITKVA 1964
QY 1915 SSPAPPEGSHPREPSTPHRY--REGRTBLRDKSPGRPLEKEKSGFGRMLSTRRRSPGR 1972
Db 1965 SSPAPPEGSHPREPSTPHRYDREGRTBLRDKSPGRPLEKEKSGFGRMLSTRRRSPGR 2024
QY 1973 LPESSRGRPLPAGAVRTPLSQVKNKRGQSA 2002
Db 2025 LPESSRGRPLPAGAVRTPLSQVKNKWDQSS 2054
```

RESULT 10

US-10-028-946-4

; Sequence 4, Application US/10028946

; Publication No. US20020123622A1

```
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Fiddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1958
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-4
```

```
Query Match 90.4%; Score 9487.5; DB 13; Length 1958;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1877; Conservative 1; Mismatches 1; Indels 81; Gaps 3;

QY 1 MLKFKYGARNPLDAGAAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFKYGARNPLDAGAAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE 60
QY 61 ECSOPALMKIKHVSNFVKYSDTIAELOELQPSAKDPEVRSLYCCGHFAEVQVVRKATG 120
Db 61 ECSOPALMKIKHVSNFVKYSDTIAELOELQPSAKDPEVRSLYCCGHFAEVQVVRKATG 120
QY 121 DIYAMKMKKALLAQOQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYOPGG 180
Db 121 DIYAMKMKKALLAQOQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYOPGG 180
QY 181 DLLSLNRYEDQDENLIQFYLAELIILAVHSVHLMGVVHRDIKPENILVDRTHGKILVDF 240
Db 181 DLLSLNRYEDQDENLIQFYLAELIILAVHSVHLMGVVHRDIKPENILVDRTHGKILVDF 240
QY 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMGDKGTGLDCDMSVGVIAEYMIYGR 300
Db 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMGDKGTGLDCDMSVGVIAEYMIYGR 300
QY 301 SPFAEGTSARTFNNIMNFQRLKFPDPPKVSDDFLDIQSLCCQKQERLKFEGLCCHPFF 360
Db 301 SPFAEGTSARTFNNIMNFQRLKFPDPPKVSDDFLDIQSLCCQKQERLKFEGLCCHPFF 360
QY 361 SKIDWNNIRSPPPFVPTLSDDDTSNFDPEPKNSWSSSPCOLSPSGFSGEELPFVGF 420
Db 361 SKIDWNNIRSPPPFVPTLSDDDTSNFDPEPKNSWSSSPCOLSPSGFSGEELPFVGF 420
QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKEQEMTRLHRRVS 480
Db 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKEQEMTRLHRRVS 480
QY 481 EVEAVLSQKEVELKASQTQSLLEQDLATYITBCSSIKRSLRLEQARMEVSEDKALQLLH 540
Db 481 EVEAVLSQKEVELKASQTQSLLEQDLATYITBCSSIKRSLRLEQARMEVSEDKALQLLH 540
QY 541 DIREOSKLOEIKBOEQYQAOVVEMLMNQLEEDLVARRRSDLYSESELRESLAAEFK 600
Db 541 DIREOSKLOEIKBOEQYQAOVVEMLMNQLEEDLVARRRSDLYSESELRESLAAEFK 600
QY 601 RKATECQKLLKAKDQKPEVGEYAKLEKINAEQQLKIQELQEKLEKA----- 648
Db 601 RKATECQKLLKAKDQKPEVGEYAKLEKINAEQQLKIQELQEKLEKAVKASTEATELLQ 660
QY 649 ----AKERARELEKLNREDSSGIRKLVAEERHSLKVRLETWERENRLKDD 704
Db 661 NIROAKERARELEKLNREDSSGIRKLVAEERHSLKVRLETWERENRLKDD 720
QY 705 IQTKSQIQQWADKILELEKHKREAQVSAQHLEVHLKQEKHQHYEKKIKVLDNQIKDLAD 764
Db 721 IQTKSQIQQWADKILELEKHKREAQVSAQHLEVHLKQEKHQHYEKKIKVLDNQIKDLAD 780
```

Qy	765	KE'YLENNVQRHEBEBAHKGKILSBQKAMINAMDSKIRSLFORIVELSEANKLAANSLSFT	824
Db	781	KET'LENNVQRHEBEBAHKGKILSBQKAMINAMDSKIRSLFORIVELSEANKLAANSLSFT	840
Qy	825	QRNKKAQEBEIMSELURQOKFYLETQAGKLEAQRNKLSEOLEKISHODSHDKRNLLLELTRL	884
Db	841	QRNKKAQEBEIMSELURQOKFYLETQAGKLEAQRNKLSEOLEKISHODSHDKRNLLLELTRL	900
Qy	885	REVSLEHEBOKLEUKROLTELQLSQLBRESQLTALQAARAALSESQLEAETTAEE	944
Db	901	REVSLEHEBOKLEUKROLTELQLSQLBRESQLTALQAARAALSESQLEAETTAEE	960
Qy	945	BEETQALTAHRDETORQFADALRNSCTVITLDEBOLNQLTBDNAELNNQNFVLSKOLDEAS	1004
Db	961	BEETQALTAHRDETORQFADALRNSCTVITLDEBOLNQLTBDNAELNNQNFVLSKOLDEAS	1020
Qy	1005	GANDEIVQLRSEVDHLRREITEREMQJTSQKQTMALKTCTMLEEQVMDLEALNDELLE	1064
Db	1021	GANDEIVQLRSEVDHLRREITEREMQJTSQKQTMALKTCTMLEEQVMDLEALNDELLE	1080
Qy	1065	KEHQWEAWRSVLGDEKQSFQFCEVRELQRMJLDTKQSRADQRTITRESSQVVELAVREHKA	1124
Db	1081	KEHQWEAWRSVLGDEKQSFQFCEVRELQRMJLDTKQSRADQRTITRESSQVVELAVREHKA	1140
Qy	1125	EIALQOALKEQKLABSLDKLNDLEKKHAMJEMNARSLOQKLETERELKQRLLEBOAK	1184
Db	1141	EIALQOALKEQKLABSLDKLNDLEKKHAMJEMNARSLOQKLETERELKQRLLEBOAK	1200
Qy	1185	LOQOMDLQKNHIFRLTQGLQEAJLDRADLLKTERSJLEYQLENTQVLYSHEKVKMEGTISQ	1244
Db	1201	LOQOMDLQKNHIFRLTQGLQEAJLDRADLLKTERSJLEYQLENTQVLYSHEKVKMEGTISQ	1260
Qy	1245	QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVPLQYNELKJALEKEKARCALKEEA	1304
Db	1261	QTKLIDFLQAKMDQPAKKKK-----VFLQYNELKJALEKEKARCALKEEA	1305
Qy	1305	LQKTRLELSARBEAAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPQSAKMLAPPSS	1364
Db	1306	LQKTRLELSARBEAAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPQSAKMLAPPSS	1365
Qy	1365	RRKESSTPEFSRRLKERMHNNIHPHFNVLGNMRAKCAVCLQTVHFGRAQSKLECOVM	1424
Db	1366	RRKESSTPEFSRRLKERMHNNIHPHFNVLGNMRAKCAVCLQTVHFGRAQSKLECOVM	1425
Qy	1425	CHPKCSTCLPATCGLPAEYATHTEAFPCRDKNMSPLQTKPEPSSSLHLEGMKVPNNKR	1484
Db	1426	CHPKCSTCLPATCGLPAEYATHTEAFPCRDKNMSPLQTKPEPSSSLHLEGMKVPNNKR	1485
Qy	1485	GOQGWDRKYIVLSEGSKVLIYDNEAREAGORPVEFELCLPDGDVSIHGAVCASELANTAK	1544
Db	1486	GOQGWDRKYIVLSEGSKVLIYDNEAREAGORPVEFELCLPDGDVSIHGAVCASELANTAK	1545
Qy	1545	A-----EKAEADAKL	1554
Db	1546	ADVPYILKXMESHPHTTCWPQRTLYLLAPSPDQKRWVTALESVVAGRVSRERKAEADAKL	1605
Qy	1555	LQNSLLKLEGGDRLDNMCTLPFSDQVVLVGTESGLYALNVKNSLTHVPGIGAVFOIYII	1614
Db	1606	LQNSLLKLEGGDRLDNMCTLPFSDQVVLVGTESGLYALNVKNSLTHVPGIGAVFOIYII	1665
Qy	1615	KDLEKLLMTAGEBRALCLVDKVKQSLAOSHUPAOPDISPNTFEAVKGCCHLFGAGKIEN	1674
Db	1666	KDLEKLLMTAGEBRALCLVDKVKQSLAOSHUPAOPDISPNTFEAVKGCCHLFGAGKIEN	1725
Qy	1675	GLCICAMPSPKVILRYNENLSKYCRKETETSEPCSCHTFTNYSLIGTNKFYEIDMKQ	1734
Db	1726	GLCICAMPSPKVILRYNENLSKYCRKETETSEPCSCHTFTNYSLIGTNKFYEIDMKQ	1785
Qy	1735	YTLBEFLDKNDHSLAPAVFAASNSFPVSIQVNSAGOREEYLLCPEHFGVFDVSYGRSS	1794
Db	1786	YTLBEFLDKNDHSLAPAVFAASNSFPVSIQVNSAGOREEYLLCPEHFGVFDVSYGRSS	1845

Qy	1795	RTDCLKWSRLPLAFAYREPYLFVTHFNLSLEVIQIARSSAGTARAYLDIPNPRYLGPAI	1854
Db	1846	RTDCLKWSRLPLAFAYREPYLFVTHFNLSLEVIQIARSSAGTARAYLDIPNPRYLGPAI	1905
Qy	1855	SSGAIYLAASSYQDKLRVICCKGNLKVESGTEHHRGPSTSR	1894
Db	1906	SSGAIYLAASSYQDKLRVICCKGNLKVESGTEHHRGPSTSR	1945

RESULT 11

US-09-964-956-40

; Sequence 40, Application US/09964956

; Publication No. US20040043926A1

; GENERAL INFORMATION:

; APPLICANT: Gerlach, Valerie L

; APPLICANT: MacDougall, John R

; APPLICANT: Smithson, Glennda

; APPLICANT: Millet, Isabelle

; APPLICANT: Stone, David

; APPLICANT: Gutther, Erik

; APPLICANT: Ellerman, Karen

; APPLICANT: Grosse, William M

; APPLICANT: Alsobrook II, John P

; APPLICANT: Lepley, Denise M

; APPLICANT: Burgess, Catherine E

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Kekkuda, Ramesh

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Leach, Martin D

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-124

; CURRENT APPLICATION NUMBER: US/09/964,956

; CURRENT FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/235,631

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/235,633

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/235,808

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,064

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,065

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,066

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,135

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: 60/237,434

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/238,321

; PRIOR FILING DATE: 2000-10-05

; PRIOR APPLICATION NUMBER: 60/238,399

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,396

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/276,667

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/294,823

; PRIOR FILING DATE: 2001-05-31

; PRIOR APPLICATION NUMBER: 60/304,868

; PRIOR FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 127

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 40

; LENGTH: 1641

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-964-956-40

QY 374 PFVPTLKSDDDTSNFDPEPKNSWSSPCQLSPSPSGBELPFVGFYSKALGILGRSES 433
Db 1 PFVPTLKSDDDTSNFDPEPKNSWSSVQCLSPSPSGBELPFVGFYSKALGILGRSES 60
QY 434 VVSGLDSPAKTSMMEKKLIIKSKELQSDQKCHKEQEMTRLHRRVSEVAVLSKEVEL 493
Db 61 VVSSLDSPAKVSSMEKKLIIKSKELQSDQKCHKEQEMTRLHRRVSEVAVLSKEVEL 120
QY 494 KASETORSLEODLATYITECSSLKRSLFQARMEVSCDDKALQLLHDIREOSRKLQBIK 553
Db 121 KASETORSLEODLATYITECSSLKRSLFQARMEVSCDDKALQLLHDIREOSRKLQBIK 180
QY 554 EOEYQACQVEEMRLMMQLBEDLVASARRSDLYESELRESRLAAEFKRAKATCQHKLAKA 613
Db 181 EOEYQACQVEEMRLMMQLBEDLVASARRSDLYESELRESRLAAEFKRAKATCQHKLAKA 240
QY 614 KOQKPEVGEYAKLEKINAEQOLKIQELQEKLEKA-----AKERAEREL 657
Db 241 KOQKPEVGEYSKLEKINAEQOLKIQELQEKLEKAVKASTEATELLQIRQAKERAEREL 300
QY 658 EKLNREDSEGIKKLVFAEERRHSLNKKVRLTMRERENRLKDDIQTKSQQIQQVAD 717
Db 301 EKLNREDSEGIKKLVFAE----- 321
QY 718 KILELEEKHREAAQVSAHLVHLKQKHQYEEKIKVLNQIKKOLADKETTLENMMORHEE 777
Db 322 ---ELEEKHREAAQVSAHLVHLKQKHQYEEKIKVLNQIKKOLADKETTLENMMORHEE 378
QY 778 EAEHKGKILSEQKAMINAMSKIRSLRIVELSEBANKLAANSSFTQRMKAQEMISE 837
Db 379 EAEHKGKILSEQKAMINAMSKIRSLRIVELSEBANKLAANSSFTQRMKAQEMISE 438
QY 838 LRQOKPYLTOAGKLEAQNKLKEOLEKISHODSDQNRLLLETRLREVSLEHEEOKLE 897
Db 439 LRQOKPYLTOAGKLEAQNKLKEOLEKISHODSDQNRLLLETRLREVSLEHEEOKLE 498
QY 898 LXRQLTELQLSQERESQTLQAARAALESQLOAKTELETTAEABEEIQALTAHRDE 957
Db 499 LXRQLTELQLSQERESQTLQAARAALESQLOAKTELETTAEABEEIQALTAHRDE 558
QY 958 IORKFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKOLDEASGANDEIVQLRSEV 1017
Db 559 IORKFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKOLDEASGANDEIVQLRSEV 618
QY 1018 DHLRREITEREMQLTSQKOTMEALKTTCMLBEOQVMDLEALNDELLEKERQWEAERSVLG 1077
Db 619 DHLRREITEREMQLTSQKOTMEALKTTCMLBEOQVMDLEALNDELLEKERQWEAERSVLG 678
QY 1078 DEKSQPECVRELOQLMDTEKQSRARADQRTESQVVELAVKEHKAETILALQOALKKEQK 1137
Db 679 DEKSQPECVRELOQLMDTEKQSRARADQRTESQVVELAVKEHKAETILALQOALKKEQK 738
QY 1138 LKAESLSDKLNLEKXHAMLENWARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIF 1197
Db 739 LKAESLSDKLNLEKXHAMLENWARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIF 798
QY 1198 RLTQGLQEAALDRADLLKTERSLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKVD 1257
Db 799 RLTQGLQEAALDRADLLKTERSLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKVD 858
QY 1258 QPAKKKKGI.FSRKEDPALPTQVPOYNELKLALEKEKARCALBEALQKTRIELRSARE 1317
Db 859 QPAKKKK-----VPOYNELKLALEKEKARCALBEALQKTRIELRSARE 903
QY 1318 EAAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPMSAMSLAPPSSRRKESSTPEEFSR 1377
Db 904 EAAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPMSAMSLAPPSSRRKESSTPEEFSR 963
QY 1378 RUKERMHHNI.PHRFNVGLNWRATKCAVCLDTHVFRQASKCLCEQVMCHPKGKSTCLPATC 1437
Db 964 RUKERMHHNI.PHRFNVGLNWRATKCAVCLDTHVFRQASKCLCEQVMCHPKGKSTCLPATC 1023

QY 1438 GLPAEYATHTEAFRCRDKMNSPGLQTKPESSSLHLEGCMKVPRNNKEQOQWDRKYIVLE 1497
Db 1024 GLPAEYATHTEAFRCRDKMNSPGLQTKPESSSLHLEGCMKVPRNNKEQOQWDRKYIVLE 1083
QY 1498 GSKVLIIDYDNEAREAGQRPVEEFELCLPDGQVSIHGAVGASELANAKA----- 1545
Db 1084 GSKVLIIDYDNEAREAGQRPVEEFELCLPDGQVSIHGAVGASELANAKADVPYILKMESH 1143
QY 1546 -----EKAEDAKLGNLSLLKLEGDDR 1567
Db 1144 HTTCWPQRTLYLLAPSPDKQWWTALFVAVAGRVSRERAEADAKLGNLSLLKLEGDDR 1203
QY 1568 LDMNCTILPFSDQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYIIKLEKLLMIAGEE 1627
Db 1204 LDMNCTILPFSDQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYIIKLEKLLMIAGEE 1263
QY 1628 RALCLVDVKKVKQSLAQSHLPAQPDISPNIFFEAVKGCFLFCAGKIENGLCICAAMPKSV 1687
Db 1264 RALCLVDVKKVKQSLAQSHLPAQPDVSPNIFFEAVKGCFLFCAGKIENSLCICAAMPKSV 1323
QY 1688 ILRYNENLSKYCIRKEIETSEPCSIHFTNYSILIGNKFEYEDMKOYTLEEFLLDKNDHS 1747
Db 1324 ILRYNENLSKYCIRKEIETSEPCSIHFTNYSILIGNKFEYEDMKOYTLEEFLLDKNDHS 1383
QY 1748 LAPAVFAASSNFPVSIQVNSAQOREEYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1807
Db 1384 LAPAVFAASSNFPVSIQVNSAQOREEYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1443
QY 1808 PAYREPVLVTFHNSLEVIHQARSSAGTAPARAYLDIPNPRYLGPALSSGAIYLIASSYQD 1867
Db 1444 PAYREPVLVTFHNSLEVIHQARSSAGTAPARAYLDIPNPRYLGPALSSGAIYLIASSYQD 1503
QY 1868 KLRVICCKGNLVKESGTEHHRGPGTSSPNKRGPPPTYNEHITKRVASSPAPPEGSPHPR 1927
Db 1504 KLRVICCKGNLVKESGTEHHRGPGTSSPNKRGPPPTYNEHITKRVASSPAPPEGSPHPR 1563
QY 1928 EPSTPHRY--REGTELRDKSGPRLERKSPGRMLSTRRERSPGRLFDSSRGRILPAG 1985
Db 1564 EPSTPHRYDRREGTELRDKSGPRLERKSPGRMLSTRRERSPGRLFDSSRGRILPAG 1623
QY 1986 AVRTPLSQVNVKGRQSA 2002
Db 1624 AVRTPLSQVNVKWDQSS 1640

RESULT 12
US-10-017-216-5
; Sequence 5, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prot
; FILE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017, 216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242, 429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1641
; TYPE: EXT
; ORGANISM: Mus musculus
US-10-017-216-5

Query Match 73.6%; Score 7717.5; DB 13; Length 1641;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 1541; Conservative 15; Mismatches 16; Indels 125; Gaps 5;

QY 374 PFVPTLKSDDDTSNFDPEPKNSWSSPCQLSPSPSGBELPFVGFYSKALGILGRSES 433
Db 1 PFVPTLKSDDDTSNFDPEPKNSWSSVQCLSPSPSGBELPFVGFYSKALGILGRSES 60

QY 434 VVSGLDSPAKTSSMEKKLLIKSKELQSDQKCHKMEQEMTRHLHRRVSEVAVLQKEVEL 493
DB 61 VVSGLDSPAKVSMKKLLIKSKELQSDQKCHKMEQEMTRHLHRRVSEVAVLQKEVEL 120
QY 494 KASETQRLSEODLATVITTECSSLSKESLEQARMEVSEQDDKALQLLHDIRQSRKLQELIK 553
DB 121 KASETQRLSEODLATVITTECSSLSKESLEQARMEVSEQDDKALQLLHDIRQSRKLQELIK 180
QY 554 EQSYQOQVEMRLMMQNLQEDLVASRRSDLYSELSRSLAAEFKQKATCOHKLKA 613
DB 181 EQSYQOQVEMRLMMQNLQEDLVASRRSDLYSELSRSLAAEFKQKATCOHKLKA 240
QY 614 KDQKPEVGEYAKLEKINAEQQLKILOELQKLEKA-----AKERAEREL 657
DB 241 KDQKPEVGEYSKLEKINAEQQLKILOELQKLEKAVKASTEATELLONTROKAREREL 300
QY 658 EKQNRDSESGIRKXKLVEABERHSLNKVRLETMERRENRLKDDIQTKSQQTQOMAD 717
DB 301 EKXHNREDSEGIKKXKLVEAE----- 321
QY 718 KILELEBKREAOVSAQHLVHLKQEQHYEKKIVLDNQIKKOLADKETLENVQRHEE 777
DB 322 ---ELEBKREAOVSAQHLVHLKQEQHYEKKIVLDNQIKKOLADKESLENVQRHEE 378
QY 778 EAHEKGIKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSLSFTQRMKAQEMISE 837
DB 379 EAHEKGIKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSLSFTQRMKAQEMISE 438
QY 838 LRQCKPYLETQAKLEAQNKLBEQLEKISHQSDXNRLLELETRLRVSLEHEBQKLE 897
DB 439 LRQCKPYLETQAKLEAQNKLBEQLEKISHQSDXNRLLELETRLRVSLEHEBQKLE 498
QY 898 LRQCKPYLETQAKLEAQNKLBEQLEKISHQSDXNRLLELETRLRVSLEHEBQKLE 957
DB 499 LRQCKPYLETQAKLEAQNKLBEQLEKISHQSDXNRLLELETRLRVSLEHEBQKLE 558
QY 958 IQKFPDALRNSCTVITDLSEQLNQLTALQAAARALLESQRLQAKTELETTABAEIEIOALTARHDE 1017
DB 559 IQKFPDALRNSCTVITDLSEQLNQLTALQAAARALLESQRLQAKTELETTABAEIEIOALTARHDE 618
QY 1018 DHRREITEREMQLTQKOTMEALKTTCTMLBEQVNDLEALNDELLEKEROEAWRSVLG 1077
DB 619 DHRREITEREMQLTQKOTMEALKTTCTMLBEQVNDLEALNDELLEKEROEAWRSVLG 678
QY 1078 DEKSQFECRVRELOQMLDTEKQSRARADQRIETESQVVELAVKEHKAETLALQALKEQK 1137
DB 679 DEKSQFECRVRELOQMLDTEKQSRARADQRIETESQVVELAVKEHKAETLALQALKEQK 738
QY 1138 LKAEISDKLNDLEKHAEMLENNARSLOQKLETERELKQLEAEQKLOQMDLQKHIF 1197
DB 739 LKAEISDKLNDLEKHAEMLENNARSLOQKLETERELKQLEAEQKLOQMDLQKHIF 798
QY 1198 RLTOGQLEALDRADLKTERTSDLEYQLENTQVLYSHKVMKMGTTISQQTKLIDFLQAKMD 1257
DB 799 RLTOGQLEALDRADLKTERTSDLEYQLENTQVLYSHKVMKMGTTISQQTKLIDFLQAKMD 858
QY 1258 QPAKKKKGLFSRRKEDPALPTQVPLQYNELKLALEKEKARCALBEALOKTRIELRSARE 1317
DB 859 QPAKKKK-----VPLQYNELKLALEKEKARCALBEALOKTRIELRSARE 903
QY 1318 EAAHRKATDHPHSTPATARQQAIAVSIRPEHQPSAMSLAPPSSRRKESSTPEFSR 1377
DB 904 EAAHRKATDHPHSTPATARQQAIAVSIRPEHQPSAMSLAPPSSRRKESSTPEFSR 963
QY 1378 RLKERMHNIPIHFNVLGNMNRATKCAVCLDTVHFGRQASKLEQCMCHPKSTCLPATC 1437
DB 964 RLKERMHNIPIHFNVLGNMNRATKCAVCLDTVHFGRQASKLEQCMCHPKSTCLPATC 1023
QY 1438 GLPAEYATHFTFAFCDDKQNSPGLQKPESSSLHLEGMKVPNNKRGQOGHDXRYIVLE 1497
DB 1024 GLPAEYATHFTFAFCDDKQNSPGLQKPESSSLHLEGMKVPNNKRGQOGHDXRYIVLE 1083

QY 1498 GSKVLIYDNEAREAGORPVEEFELCLPDGDSVSHGAVGASELANAKA----- 1545
DB 1084 GSKVLIYDNEAREAGORPVEEFELCLPDGDSVSHGAVGASELANAKADVPYILKMESH 1143
QY 1546 -----EKAEDAKLLGNSLLKLEGDDR 1567
DB 1144 HTTCWPGRGRTYLLAPSPFDKQRMWVTALESVVAGGRVSRKAEADAKLLGNSLLKLEGDDR 1203
QY 1568 LDMNCTIPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYIITKOLEKLMIAAGEE 1627
DB 1204 LDMNCTIPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYIITKOLEKLMIAAGEE 1263
QY 1628 RALCLAVDVKKVKQSLAQSHLPAQPDISPNIFFAVKGCGLHFGAGKIENGLCICAAMPKVV 1687
DB 1264 RALCLAVDVKKVKQSLAQSHLPAQPDISPNIFFAVKGCGLHFGAGKIENGLCICAAMPKVV 1323
QY 1688 ILRYNENLSKYCIKRIKTEITSEPCSIHFTNYSILITGNKFIYEDMKQYTLLEFELDKNDHS 1747
DB 1324 ILRYNENLSKYCIKRIKTEITSEPCSIHFTNYSILITGNKFIYEDMKQYTLLEFELDKNDHS 1383
QY 1748 LAPAVFAASSNSFPVSIQVNSAGQREYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1807
DB 1384 LAPAVFAASSNSFPVSIQVNSAGQREYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1443
QY 1808 FAYREPYLVFTHFNSLEVIQIARSSAGTTPARAYLDIPNPRYLGPALSSGAIYLIASSYQD 1867
DB 1444 FAYREPYLVFTHFNSLEVIQIARSSAGTTPARAYLDIPNPRYLGPALSSGAIYLIASSYQD 1503
QY 1868 KLRVICCKGNLVKESGTEHHRGPGSTSPSSPKGPPPTNYNEHITKRVASSPAPPEGSHPR 1927
DB 1504 KLRVICCKGNLVKESGTEHHRGPGSTSPSSPKGPPPTNYNEHITKRVASSPAPPEGSHPR 1563
QY 1928 EPSTPHRY--REGTELRDRDKSPGRPLERKSPGRMLSTRERSPGRLFDSSRGRPLAG 1985
DB 1564 EPSTPHRYDRREGTELRDRDKSPGRPLERKSPGRMLSTRERSPGRLFDSSRGRPLAG 1623
QY 1986 AVRTPLSOVNVKGRQSA 2002
DB 1624 AVRTPLSOVNVKGRQSS 1640

RESULT 13

US-09-964-956-41
; Sequence 41, Application US/09964956
; Publication No. US20040043926A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964,956
; PRIOR FILING DATE: 2001-09-26
; PRIOR FILING DATE: 2000-09-27
; PRIOR FILING DATE: 2000-09-27
; PRIOR FILING DATE: 2000-09-27
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR APPLICATION NUMBER: 60/236,064

; PRIOR FILING DATE: 2000-09-27									
; PRIOR APPLICATION NUMBER: 60/236,065									
; PRIOR FILING DATE: 2000-09-27									
; PRIOR APPLICATION NUMBER: 60/236,066									
; PRIOR FILING DATE: 2000-09-27									
; PRIOR APPLICATION NUMBER: 60/236,135									
; PRIOR FILING DATE: 2000-09-28									
; PRIOR APPLICATION NUMBER: 60/237,434									
; PRIOR FILING DATE: 2000-10-03									
; PRIOR APPLICATION NUMBER: 60/238,321									
; PRIOR FILING DATE: 2000-10-05									
; PRIOR APPLICATION NUMBER: 60/238,399									
; PRIOR FILING DATE: 2000-10-06									
; PRIOR APPLICATION NUMBER: 60/238,396									
; PRIOR FILING DATE: 2000-10-06									
; PRIOR APPLICATION NUMBER: 60/276,667									
; PRIOR FILING DATE: 2001-03-16									
; PRIOR APPLICATION NUMBER: 60/294,823									
; PRIOR FILING DATE: 2001-05-31									
; PRIOR APPLICATION NUMBER: 60/304,868									
; PRIOR FILING DATE: 2001-07-12									
; NUMBER OF SEQ ID NOS: 127									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 41									
; LENGTH: 1597									
; TYPE: PRT									
; ORGANISM: Mus musculus									
US-09-964-956-41									
Query Match									
Best Local Similarity 71.4%; Score 7491.5; DB 12; Length 1597;									
Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;									
Qy	468	MQEMTRLHRRVSEVAVLSQSEVELKASSETQSRSLLEDQLATYITECSSLSKRSLEQARME	527						
Db	9	MQEMTRLHRRVSEVAVLSQSEVELKASSETQSRSLLEDQLATYITECSSLSKRSLEQARME	68						
Qy	528	VSQEDDKALCLLHDIREQSRKLOEIKQEQYQAVSEMLMNMNQLBEDLVSARRRSDLYES	587						
Db	69	VSQEDDKALCLLHDIREQSRKLOEIKQEQYQAVSEMLMNMNQLBEDLVSARRRSDLYES	128						
Qy	588	ELRESLAAEFKPKKATECOHKLKAKQOGKEVGEYAKLEKINAEQQLKIQELOEKLEK	647						
Db	129	ELRESLAAEFKPKKANEQCHLMKAKQOGKEVGEYSKLEKINAEQQLKIQELOEKLEK	188						
Qy	648	A-----AKERAERELEKLNREDSESGIRKKLVABERRHSLKNVKRL	691						
Db	189	AVKASTEATELLQNIQAKERAERELEKLNREDSESGIKKLVABERRHSLKNVKRL	248						
Qy	692	ETWERRENRLKDDIOTKSOIQOMADKILEBEKEKEACVSAOHLVHLKQEQHYBEKI	751						
Db	249	ETWERRENRLKDDIOTKSEQIQOMADKILEBEKEKEACVSAOHLVHLKQEQHYBEKI	308						
Qy	752	KVLNDQIKKDLADKETTLENMQRHEEFAHEKGIILSEQKAMINAMDSKIRSLQRIVELS	811						
Db	309	KVLNDQIKKDLADKESLENMQRHEEFAHEKGIILSEQKAMINAMDSKIRSLQRIVELS	368						
Qy	812	EANKLAANSSLTQNRNKAQEMISELROOKPYLETQACKLEAQRNKLEEQLEKISHODH	871						
Db	369	EANKLAANSSLTQNRNKAQEMISELROOKPYLETQACKLEAQRNKLEEQLEKISHODH	428						
Qy	872	SDKNRLELETRLREVSLEHEKQKLELQKLTQLSLQERESQLTALQAARAALESQRL	931						
Db	429	SDKSRLELETRLREVSLEHEKQKLELQKLTQLSLQERESQLTALQAARAALESQRL	488						
Qy	932	QAKTELEETTAABEBIQALTARDBIQKFPALRNSCTVIITDLEEQNLQNTEDNAELNN	991						
Db	489	QAKTELEETTAABEBIQALTARDBIQKFPALRNSCTVIITDLEEQNLQNTEDNAELNN	548						
Qy	992	QNFYLSKQLDDEASGANDEIVQLRSEVDHLRRRITITERMOLTSQKOTMEALKTTTCMLBEQ	1051						
Db	549	QNFYLSKQLDDEASGANDEIVQLRSEVDHLRRRITITERMOLTSQKOTMEALKTTTCMLBEQ	608						

Qy	1052	VMDLEALNDELLEKERQWEAWRSVLGDEKSFQECRVRELQRLMDTEKQSRARADORITES	1111						
Db	609	VLDLEALNDELLEKERQWEAWRSVLGDEKSFQECRVRELQRLMDTEKQSRARADORITES	668						
Qy	1112	QVVVELAVKHEKAEIILALQOALKEOKLKAESLSDKNDLEKKHAMLEMNARSLOOKLETE	1171						
Db	669	QVVVELAVKHEKAEIILALQOALKEOKLKAESLSDKNDLEKKHAMLEMNARSLOOKLETE	728						
Qy	1172	RELKORLLEEQAKLQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYQLENIQVLY	1231						
Db	729	RELKORLLEEQAKLQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYQLENIQVLY	788						
Qy	1232	SHEKVKMEGTISOOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVFLQYNELKAL	1291						
Db	789	SHEKVKMEGTISOOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVFLQYNELKAL	833						
Qy	1292	EKEKARCAELEALOKTRIELRSAREEAAHRKATDHPHPSTPATARQIAMSALVRSPEH	1351						
Db	834	EKEKARCAELEALOKTRIELRSAREEAAHRKATDHPHPSTPATARQIAMSALVRSPEH	893						
Qy	1352	QPSAMSLAPPSRRRKESSTPEFSRRLKERMHNIPHRNVGLNMRATKCAVCLDVTVFH	1411						
Db	894	QPSAMSLAPPSRRRKESSTPEFSRRLKERMHNIPHRNVGLNMRATKCAVCLDVTVFH	953						
Qy	1412	GRQASKCLEQVMCHPKCSTCLPATCGLPAEYATHETEARFCRDKMNSPGLQTEPSSSLH	1471						
Db	954	GRQASKCLEQVMCHPKCSTCLPATCGLPAEYATHETEARFCRDKMNSPGLQTEPSSSLH	1013						
Qy	1472	LEGWMKVPNNKRGQOGWDRKIYVLEGSKVLIVDNEAREAGQRPVEFECLPDGDVSIH	1531						
Db	1014	LEGWMKVPNNKRGQOGWDRKIYVLEGSKVLIVDNEAREAGQRPVEFECLPDGDVSIH	1073						
Qy	1532	GAVGASLANTAKA-----	1545						
Db	1074	GAVGASLANTAKADVPIILKMESHPTTCWPGRTLYLLAPSPDKQKRWVTALESVVVAGG	1133						
Qy	1546	---EKAEDAALLGNLSLLLEGDDRLDMNCTLPFSQOVVLVGTGEGLYALNVLKNSLTH	1601						
Db	1134	RVSREKAEAAKLLGNLSLLLEGDDRLDMNCTLPFSQOVVLVGTGEGLYALNVLKNSLTH	1193						
Qy	1602	VPGIGAVFYIYIKOLEKLLMIAGERALCLVDVKKVKQSLAOSHLPAPQDPISNIFEAV	1661						
Db	1194	IPGIGAVFYIYIKOLEKLLMIAGERALCLVDVKKVKQSLAOSHLPAPQDPISNIFEAV	1253						
Qy	1662	KGCHLFAGAKIENCLCICAAMPSKVILRYNENLSKYCIRKEIETSPSCSIHFTNYSIL	1721						
Db	1254	KGCHLFAGAKIENCLCICAAMPSKVILRYNENLSKYCIRKEIETSPSCSIHFTNYSIL	1313						
Qy	1722	IGTNKFYEIDMKQYTLBEFLDKNDHSLAPAVFAASSNSFFVSIQVNSAGOREEYLLCFH	1781						
Db	1314	IGTNKFYEIDMKQYTLBEFLDKNDHSLAPAVFAASSNSFFVSIQVNSAGOREEYLLCFH	1373						
Qy	1782	EFQVFDVSYGRSRTDDLKWSRLPLAPAYREPVLVTHFNSLEVIEIQARSSAGTPARAY	1841						
Db	1374	EFQVFDVSYGRSRTDDLKWSRLPLAPAYREPVLVTHFNSLEVIEIQARSSAGTPARAY	1433						
Qy	1842	LDIENPYLGPALSSSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGPTSRSSPNKRG	1901						
Db	1434	LEIENPYLGPALSSSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGPTSRSSPNKRG	1493						
Qy	1902	PPTYNEHITKRVASSPAPGSPHREPSTPHRY--REGTELRDRKSPGRPLEREKSPG	1959						
Db	1494	PPTYNEHITKRVASSPAPGSPHREPSTPHRY--REGTELRDRKSPGRPLEREKSPG	1553						
Qy	1960	RMLSTRERSPGRLPEDSSGRPLPAGAVRTPLSQVNNKRGQSA	2002						
Db	1554	RMLSTRERSPGRLPEDSSGRPLPAGAVRTPLSQVNNKWDQSS	1596						

RESULT 14
US-10-017-216-6
; Sequence 6, Application US/10017216
; Publication No. US20020160483A1


```

; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prob
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-6

Query Match 71.4%; Score 7491.5; DB 13; Length 1597;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;

Qy 468 MEQEMTLRHRVSEVAVLSQKEVELKASQTSQSLLEQDLATVITCSSLKRSLEQARME 527
Db 9 MEQEMTLRHRVSEVAVLSQKEVELKASQTSQSLLEQDLATVITCSSLKRSLEQARME 68

Qy 528 VSQEDDKALQLLHDHDEQSRKLEIIEQOEYQAOVEEMRLMMNQLBEDLVSARRSDLYES 587
Db 69 VSQEDDKALQLLHDHDEQSRKLEIIEQOEYQAOVEEMRLMMNQLBEDLVSARRSDLYES 128

Qy 588 ELRESLAAEFKFKKATECOHLLKAKDQKPGVGYAKLEKINAEQQLKIQBLQEKLEK 647
Db 129 ELRESLAAEFKFKKANECHOLMKAKDQKPGVGYKLEKINAEQQLKIQBLQEKLEK 188

Qy 648 A-----AKERARELEKIQNRDSSSEGIKKLVEAERHRSLENKVKRL 691
Db 189 AVKASTEATELLQNRQAKERARELEKLNHRDSSSEGIKKLVEAERHRSLENKVKRL 248

Qy 692 ETWERENRLKDDIQTKSQIQQMAQKILELEKREHQAQVSAQHLEVHLKQEQHYEKKI 751
Db 249 ETWERENRLKDDIQTKSQIQQMAQKILELEKREHQAQVSAQHLEVHLKQEQHYEKKI 308

Qy 752 KVLDDQIKKOLADKETLENMQRHEPEAEKGIKLEQKAMINAMDSKIRSLQRIVELS 811
Db 309 KVLDDQIKKOLADKESLENMQRHEPEAEKGIKLEQKAMINAMDSKIRSLQRIVELS 368

Qy 812 EANKLAANSLSLFTQNNKAOEEMISLRQOKFVLETOAGKLEAQNKLBEQLEKISHQDH 871
Db 369 EANKLAANSLSLFTQNNKAOEEMISLRQOKFVLETOAGKLEAQNKLBEQLEKISHQDH 428

Qy 872 SDNRILLELETRLREVSLEHEEQKLEKQLTELQLSQERESQLTALQAAARAALESQIR 931
Db 429 SDKSRILLELETRLREVSLEHEEQKLEKQLTELQLSQERESQLTALQAAARAALESQIR 488

Qy 932 QAKTELEETTAABESIQALTAHRDIBIQKFDALRNSCTVITDLBEQLNQLTEDNAELNN 991
Db 489 QAKTELEETTAABESIQALTAHRDIBIQKFDALRNSCTVITDLBEQLNQLTEDNAELNN 548

Qy 992 QNFYLSKQLDEASGANDETVQLRSEVDHLRREITEREMOLTSQKOTMEALKTTCTMLBEQ 1051
Db 549 QNFYLSKQLDEASGANDETVQLRSEVDHLRREITEREMOLTSQKOTMEALKTTCTMLBEQ 608

Qy 1052 VMDEALNDELLEKPERQEWANRSLVDEKQSECFRVELRQRLMDTEKSRARADQRITES 1111
Db 609 VLDLEALNDELLEKPERQEWANRSLVDEKQSECFRVELRQRLMDTEKSRARADQRITES 668

Qy 1112 RQVVELAVKHEKABEIALQALKEQKLEKAEESLSDKLNLEKHAMLENNARSLQOKLETE 1171
Db 669 RQVVELAVKHEKABEIALQALKEQKLEKAEESLSDKLNLEKHAMLENNARSLQOKLETE 728

Qy 1172 RELKORLLEEQAKLQOQMDLQKNHIFRLTQGLQEALDRADLKLKTERSDLEYQLENIQVLY 1231
Db 729 RELKORLLEEQAKLQOQMDLQKNHIFRLTQGLQEALDRADLKLKTERSDLEYQLENIQVLY 788

```

```

Qy 1232 SHEKVKMEGTISQOTKLIDFLQAKMDQPAKCKKGLFSRRKEDPALPTQVPLQYNELKAL 1291
Db 789 SHEKVKMEGTISQOTKLIDFLQAKMDQPAKCKK-----VPLQYNELKAL 833

Qy 1292 EKEKARCAELAEALOKTRIELRSAREEAHRKATDHPSTPATARQOIAMSIVSPH 1351
Db 834 EKEKARCAELAEALOKTRIELRSAREEAHRKATDHPSTPATARQOIAMSIVSPH 893

Qy 1352 QPSAMSLIAPPSSRRKESSTPEEFSRRLKERMHNIPIHFRNVGLNMEATKCAVCLDTVHF 1411
Db 894 QPSAMSLIAPPSSRRKESSTPEEFSRRLKERMHNIPIHFRNVGLNMEATKCAVCLDTVHF 953

Qy 1412 GRQASKLCLEQVMCHPKCSTCLPATCGLPAEYATHFTAEFCRDKNMSPGLOKPESSSLH 1471
Db 954 GRQASKLCLEQVMCHPKCSTCLPATCGLPAEYATHFTAEFCRDKNMSPGLOKPESSSLH 1013

Qy 1472 LEGNMKVPNNKRGQOGWDRKYIVLEGSKVLIIYDNEAREAGQRPVEFELCLPDGDVSIH 1531
Db 1014 LEGNMKVPNNKRGQOGWDRKYIVLEGSKVLIIYDNEAREAGQRPVEFELCLPDGDVSIH 1073

Qy 1532 GAVGASELANATAK----- 1545
Db 1074 GAVGASELANATAKADVPYILKMESHPHHTTCWPGRTLYLLAPSPDKQRWVTALESVVAGG 1133

Qy 1546 ----EKAEDAADAKLLGNSLLKLEGGDRLDMNCTLPFSDQVVLVGTTEGLYALNVLKNSLTH 1601
Db 1134 RVSREKAEDAADAKLLGNSLLKLEGGDRLDMNCTLPFSDQVVLVGTTEGLYALNVLKNSLTH 1193

Qy 1602 VPGIGAVFQYIIVKLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAV 1661
Db 1194 IPGIGAVFQYIIVKLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDVSPN-PEAV 1253

Qy 1662 KGCHLFGAGKIENGLCICAAMPKSVILRYNENLSKYCIKEIETSEPCSHFTNYSIL 1721
Db 1254 KGCHLFGAGKIENGLCICAAMPKSVILRYNENLSKYCIKEIETSEPCSHFTNYSIL 1313

Qy 1722 IGTKNFYEIDMKQYTLLEFLLDKNDHSLAPAVFAASSNSFPVSIQVNSAQOREEYLLCFH 1781
Db 1314 IGTKNFYEIDMKQYTLLEFLLDKNDHSLAPAVFAASSNSFPVSIQVNSAQOREEYLLCFH 1373

Qy 1782 EFGVVFDSYGRSRDLDLKWRLPLAFAYREPILFVTHFNSLEVIQARSSAGTPARAY 1841
Db 1374 EFGVVFDSYGRSRDLDLKWRLPLAFAYREPILFVTHFNSLEVIQARSSAGTPARAY 1433

Qy 1842 LDIENPRYLPAPASSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPSTSSSPNKR 1901
Db 1434 LEIENPRYLPAPASSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPSTSSSPNKR 1493

Qy 1902 PPTVNEHITKRVASSPAPPEGSHPREPSTPHRY--REGTELRDRKSPGPRERKSPG 1959
Db 1494 PPTVNEHITKRVASSPAPPEGSHPREPSTPHRYDRREGTELRDRKSPGPRERKSPG 1553

Qy 1960 RMLSTRERSGRLFEDESSRGRLPAGAVRTPLSQVNNKRGOSA 2002
Db 1554 RMLSTRERSGRLFEDESSRGRLPAGAVRTPLSQVNNKRGOSA 1596

RESULT 15
US-10-412-897-3
; Sequence 3, Application US/10412897
; Publication No. US2003020224A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES ENCODING THE HUMAN CITRON KINASE
; TITLE OF INVENTION: POLYPEPTIDE, BMSNKC_0020/0021
; FILE REFERENCE: D0193 NP
; CURRENT APPLICATION NUMBER: US/10/412,897
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: U.S. 60/372,745
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3

```

1518 EFELCLPDGDSVHGAASELANTAKA-----1545
1546 -----EKAADAKLLGNSLLKLEGGDDRLDMNCTLPFSDQVVLVGTTEE 1587
1588 GLYALNVKNSLTHVPGIGAVFQYIIKDKLEKLLMTAGEERALCLVDVKKVQSLAOSHL 1647
1648 PAQDIPSPNIFEAVKGCHELFCAKGIENGLCICAMPSSKVILRYNENLSKYCKEIEETS 1707
1708 BPCSCIHFTNYSILIGTNKFEYIDMKOYTLLEBFLDKNDHSLAPAVFAASSNSPVSIVQV 1767
1768 NSAQOREEYLLCFHEFGVFDVSYGRRSRTDLDLKWRLPLAFAYREPFLFVTHFNSLEVIE 1827
1828 IOARSSAGTARAYLDIPNPRYLGPALSSGAIYLAASYQDKLAVICCKGNLVKESGTEHH 1887
1888 RGPSTSSSPNKGPPTYNESHITKRVASSPAPPEGSPHREPSTPHRY--REGTELRDD 1945
1946 KSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRGQSA 2002
2002 KSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRGQSA 2002
2002 KSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRGQSA 2002

Search completed: July 3, 2004, 10:26:03
Job time : 110 secs

QY 1518 EFELCLPDGDSVHGAASELANTAKA-----1545
Db 1546 -----EKAADAKLLGNSLLKLEGGDDRLDMNCTLPFSDQVVLVGTTEE 1587
QY 1588 GLYALNVKNSLTHVPGIGAVFQYIIKDKLEKLLMTAGEERALCLVDVKKVQSLAOSHL 1647
Db 1648 PAQDIPSPNIFEAVKGCHELFCAKGIENGLCICAMPSSKVILRYNENLSKYCKEIEETS 1707
QY 1708 BPCSCIHFTNYSILIGTNKFEYIDMKOYTLLEBFLDKNDHSLAPAVFAASSNSPVSIVQV 1767
Db 1768 NSAQOREEYLLCFHEFGVFDVSYGRRSRTDLDLKWRLPLAFAYREPFLFVTHFNSLEVIE 1827
QY 1828 IOARSSAGTARAYLDIPNPRYLGPALSSGAIYLAASYQDKLAVICCKGNLVKESGTEHH 1887
Db 1888 RGPSTSSSPNKGPPTYNESHITKRVASSPAPPEGSPHREPSTPHRY--REGTELRDD 1945
QY 1946 KSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRGQSA 2002
Db 2002 KSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRGQSA 2002
2002 KSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRGQSA 2002

Query Match 64.3%; Score 6745.5; DB 15; Length 1441;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 1345; Conservative 15; Mismatches 12; Indels 125; Gaps 5;

QY 574 DLVSARRRSDLYSELRESRLAAEFKPKATCQHKLLKAKDOGKPEVGEYAKLEKINAE 633
Db 1 DLVSARRRSDLYSELRESRLAAEFKPKATCQHKLLKAKDOGKPEVGEYAKLEKINAE 60
QY 634 QOLKIQLOEKLEKA-----AKERAERELEKLNKREDSSEGIKKLVEA 677
Db 61 QOLKIQLOEKLEKAVKASTEALLQNIQAKERAERELEKLNKREDSSEGIKKLVEA 120
QY 678 EBRHSLNKVRLETMERENRLKDDIQTKSQIQOMADKILEBEKHREAOVSAQHLE 737
Db 121 E-----ELEKHREAOVSAQHLE 138
QY 738 VHLKOKEQHYEBKIKVLDNQIKDLADKETENMMQORHEEBAHEKGIKILSEKAMINAMD 797
Db 139 VHLKOKEQHYEBKIKVLDNQIKDLADKESLENNQORHEEBAHEKGIKILSEKAMINAMD 198
QY 798 SKIRSLQRIVELSEANKLAANSSFTQRMKAQEEI-SELQKQFYLETQAGKLEAQRN 857
Db 199 SKIRSLQRIVELSEANKLAANSSFTQRMKAQEEI-SELQKQFYLETQAGKLEAQRN 258
QY 858 KLEBQLEKISHODSDKNRLLEFLRLRVSLEHEFOKLEKXOLTELQSLQERESOLT 917
Db 259 KLEBQLEKISHODSDKSLLEFLRLRVSLEHEFOKLEKXOLTELQSLQERESOLT 318
QY 918 ALQARAALLESQLOAKTELETTAEAEIEIQAHTAHREIQRKFDALRNSCTVITDLEE 977
Db 319 ALQARAALLESQLOAKTELETTAEAEIEIQAHTAHREIQRKFDALRNSCTVITDLEE 378
QY 978 QLNQLTEDNAELNNQNYLSKQLDASGANDIIVQIRSEVDHLRREITERENQLTQSQKT 1037
Db 379 QLNQLTEDNAELNNQNYLSKQLDASGANDIIVQIRSEVDHLRREITERENQLTQSQKT 438
QY 1038 MEALTKTCTMLEEQVMDLEALDELLEKEROEAWRSVLGDEKSFECRVRELQRLMDTE 1097
Db 439 MEALTKTCTMLEEQVMDLEALDELLEKEROEAWRSVLGDEKSFECRVRELQRLMDTE 498
QY 1098 KOSRARADQRIITESRQVVELAVKEHKAETILALQALKEQKKAESLSDKINDLEKXHML 1157
Db 499 KOSRARADQRIITESRQVVELAVKEHKAETILALQALKEQKKAESLSDKINDLEKXHML 558
QY 1158 ENNARSLOQKLETERELKORLLEEQAKLOQOQMDLQKNHIFRLTQGLQALDRADLLKTER 1217
Db 559 ENNARSLOQKLETERELKORLLEEQAKLOQOQMDLQKNHIFRLTQGLQALDRADLLKTER 618
QY 1218 SDLEYQLENIQVLYSHEKVMKSGTISQQTKLIDFLQAKMDQPAKKKGLFRRKEDPALP 1277
Db 619 SDLEYQLENIQVLYSHEKVMKSGTISQQTKLIDFLQAKMDQPAKKKGLFRRKEDPALP 665
QY 1278 TOVPOYNEELKALKEKARCAELBEALQKTRIELRSAREEAAHRAKATDHPHPSTPATAR 1337
Db 666 --VPLQYNEELKALKEKARCAELBEALQKTRIELRSAREEAAHRAKATDHPHPSTPATAR 723
QY 1338 QOIAMSALVRSPEHOPASMSLLAPPSSRREKSTPEEFSSRLKERMHNIPIHFRNVGLNM 1397
Db 724 QOIAMSALVRSPEHOPASMSLLAPPSSRREKSTPEEFSSRLKERMHNIPIHFRNVGLNM 783
QY 1398 RATKCAVCLDTHVHGRQASKCLEQVMCHPKCTCLPATCGLPAEYATHFTFAFCRDXN 1457
Db 784 RATKCAVCLDTHVHGRQASKCLEQVMCHPKCTCLPATCGLPAEYATHFTFAFCRDXN 843
QY 1458 SPGLQTKPESSSLHLEGMKVPNNKRGQOQNDKRYIVLEGSKVLIYDNEAREAGQRPVE 1517
Db 844 SPGLQTKPESSSLHLEGMKVPNNKRGQOQNDKRYIVLEGSKVLIYDNEAREAGQRPVE 903

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2004, 09:41:14 ; Search time 2580 Seconds
(without alignments)
12292.380 Million cell updates/sec

Title: US-10-017-216-1

Perfect score: 6574

Sequence: 1 agagccgcagtgaggagat.....atcgagaatgtaggttaga 6574

Scoring table:

IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6574	100.0	6574	14	US-10-017-216-1
2	6574	100.0	6574	15	US-10-325-430-10
3	6162	93.7	6162	15	US-10-325-430-11
4	6159	93.7	6159	14	US-10-017-216-3
5	5684.4	86.5	6298	13	US-10-415-011-43
6	5661.4	86.1	6165	14	US-10-028-946-1
7	5631	85.7	6189	13	US-10-282-511-1
8	5631	85.7	6189	13	US-09-964-956-10
9	5629.8	85.6	6201	13	US-10-262-511-13
10	5629.8	85.6	6201	13	US-09-964-956-8
11	5358	81.5	5877	14	US-10-028-946-3
12	2564.8	39.0	3131	13	US-10-276-774-137
13	2414	36.7	2542	13	US-10-282-511-7
14	2317.2	35.2	2497	13	US-10-282-511-5

15	1841.2	28.0	1870	13	US-10-262-511-3	Sequence 3, Appli
16	1524.2	23.2	1870	13	US-10-262-511-9	Sequence 9, Appli
17	1467.6	22.3	1915	13	US-10-262-511-11	Sequence 11, Appli
18	1406.4	21.4	2066	17	US-10-311-034-33	Sequence 33, Appli
19	1397.8	21.3	1515	9	US-09-804-471A-1	Sequence 1, Appli
20	1397.8	21.3	1515	15	US-10-238-709-1	Sequence 1, Appli
21	1397.8	21.3	1515	17	US-10-724-594-1	Sequence 1, Appli
22	965.6	14.7	2162	16	US-10-120-988-419	Sequence 419, App
23	950.6	14.5	957	16	US-10-412-897-1	Sequence 1, Appli
24	768	11.7	1133	15	US-10-282-048-1	Sequence 1, Appli
25	701.4	10.7	995	13	US-10-276-774-255	Sequence 255, App
26	421.4	6.4	2870	13	US-10-425-114-26241	Sequence 26241, A
27	350.4	5.3	354	14	US-10-040-739-9	Sequence 9, Appli
28	308	4.7	308	9	US-09-998-598-2328	Sequence 2328, Ap
29	306.4	4.7	308	9	US-09-815-343-132	Sequence 132, App
30	306.4	4.7	308	9	US-09-998-598-1929	Sequence 1929, Ap
31	306.4	4.7	308	13	US-10-097-105-132	Sequence 132, App
32	296	4.5	297	9	US-09-796-632-3885	Sequence 3885, Ap
33	296	4.5	297	15	US-10-040-862-3885	Sequence 3885, Ap
34	296	4.5	297	16	US-10-057-475B-3885	Sequence 3885, Ap
35	296	4.5	297	16	US-10-154-884B-3885	Sequence 3885, Ap
36	258	3.9	258	16	US-10-305-720-513	Sequence 513, App
37	243.6	3.7	1530	13	US-10-415-011-24	Sequence 24, Appli
38	243.6	3.7	4698	17	US-10-702-496-1	Sequence 1, Appli
39	243.6	3.7	5373	13	US-10-333-314-40	Sequence 40, Appli
40	236.8	3.6	6335	16	US-10-388-934-85	Sequence 85, Appli
41	234	3.6	5373	17	US-10-702-496-5	Sequence 5, Appli
42	228.2	3.5	2785	13	US-10-342-687-930	Sequence 930, App
43	228.2	3.5	2785	13	US-10-172-118-930	Sequence 930, App
44	228.2	3.5	5694	13	US-10-362-892-26	Sequence 26, Appli
45	228.2	3.5	5694	16	US-10-288-798-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1

US-10-017-216-1
; Sequence 1, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPILLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1 Human Myotonic Dystrophy Type Prote
; TITLE OF INVENTION: Kinase and Uses Thereof
; FILE REFERENCE: 10147-57UI
; CURRENT APPLICATION NUMBER: US/10/017,216
; PRIOR FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-216-1

Query Match	100.0%	Score 6574;	DB 14;	Length 6574;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 6574;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCT	60	
Db	1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCT	60	
Qy	61	GGTGCTGTGAACCCATTGCGAGCGGGCTCCAGCTGAATCTGTTCTCCAGGGGAAA	120	
Db	61	GGTGCTGTGAACCCATTGCGAGCGGGCTCCAGCTGAATCTGTTCTCCAGGGGAAA	120	
Qy	121	CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGGAAGGATATTAGATGCC	180	
Db	121	CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGGAAGGATATTAGATGCC	180	

QY 4561 CCGGTGGAAGAAATTTGAGCTGTGCTTCCCGAGGGAGTGTATCTATTATTCATGGTCCCGTT 4620
Db 4561 CCGGTGGAAGAAATTTGAGCTGTGCTTCCCGAGGGAGTGTATCTATTATTCATGGTCCCGTT 4620
QY 4621 GGTGCTTCCGAATCTCGAAATACAGCCAAAGCAGAAAAGCAGAAAGCTGATCTAAACTG 4680
Db 4621 GGTGCTTCCGAATCTCGAAATACAGCCAAAGCAGAAAAGCAGAAAGCTGATCTAAACTG 4680
QY 4681 CTTTGGAAATCTCCCTGCTGAAACTGGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4740
Db 4681 CTTTGGAAATCTCCCTGCTGAAACTGGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4740
QY 4741 CCCTTCAGTGACAGAGTGTGTTGGTGGGACCGAGGAAGGGCTCTACGCCCTGAAATGTC 4800
Db 4741 CCCTTCAGTGACAGAGTGTGTTGGTGGGACCGAGGAAGGGCTCTACGCCCTGAAATGTC 4800
QY 4801 TTGAAAAATCTCCCTAACCATGTCCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
Db 4801 TTGAAAAATCTCCCTAACCATGTCCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
QY 4861 AAGGACTGGAGAGTACTCATATGATAGCAGGAAGAGCGGGCACTGTGTCTTGTGGAC 4920
Db 4861 AAGGACTGGAGAGTACTCATATGATAGCAGGAAGAGCGGGCACTGTGTCTTGTGGAC 4920
QY 4921 GTCAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCCCTGCCAGCCGACATCTCA 4980
Db 4921 GTCAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCCCTGCCAGCCGACATCTCA 4980
QY 4981 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5040
Db 4981 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5040
QY 5041 GGGCTCTGCAATCTGTGCGAGGCAATGCCAGAAAGTGTCAATCTCCGCTACAAAGAAAC 5100
Db 5041 GGGCTCTGCAATCTGTGCGAGGCAATGCCAGAAAGTGTCAATCTCCGCTACAAAGAAAC 5100
QY 5101 CTGAGCAAAATCTGCAATCTCGGAAAGAGATAGAGACCTCAGAGCCCTGCGAGCTGTATCCAC 5160
Db 5101 CTGAGCAAAATCTGCAATCTCGGAAAGAGATAGAGACCTCAGAGCCCTGCGAGCTGTATCCAC 5160
QY 5161 TTCAACAAATPACAGTATCTCTATGGAACCAATAAATTTCTAGAAATCGACATGAAGCAG 5220
Db 5161 TTCAACAAATPACAGTATCTCTATGGAACCAATAAATTTCTAGAAATCGACATGAAGCAG 5220
QY 5221 TACACGCTCGAGAAATCTGGAATAGAAATGACCATTCCTTGGCACTGCTGTGTTGCC 5280
Db 5221 TACACGCTCGAGAAATCTGGAATAGAAATGACCATTCCTTGGCACTGCTGTGTTGCC 5280
QY 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAAACAGCGCAGGCGAGGAG 5340
Db 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAAACAGCGCAGGCGAGGAG 5340
QY 5341 GAGTACTGCTGTGTTTCCAGAAATTTGGAGTGTTCGTTGATTTCTTACGGAAAGCTGAGC 5400
Db 5341 GAGTACTGCTGTGTTTCCAGAAATTTGGAGTGTTCGTTGATTTCTTACGGAAAGCTGAGC 5400
QY 5401 CGCACAGACGATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGCCCTACAGAGAACCCTAT 5460
Db 5401 CGCACAGACGATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGCCCTACAGAGAACCCTAT 5460
QY 5461 CTGTTTGTGACCACTTCAACTCAGTCTGAAAGTAATTTGAGATCCAGGCAGGCTCCTCAGCA 5520
Db 5461 CTGTTTGTGACCACTTCAACTCAGTCTGAAAGTAATTTGAGATCCAGGCAGGCTCCTCAGCA 5520
QY 5521 GGGACCCCTCGCCGAGGCTACCTGGACATCCCGGACCTCGCTACCTGGGCCCTGCCATT 5580
Db 5521 GGGACCCCTCGCCGAGGCTACCTGGACATCCCGGACCTCGCTACCTGGGCCCTGCCATT 5580
QY 5581 TCCTCAGGAGCGAATTTCTTGGGCTCCTCATACAGGATAAATTAAGGTGCTATTTGCTGC 5640
Db 5581 TCCTCAGGAGCGAATTTCTTGGGCTCCTCATACAGGATAAATTAAGGTGCTATTTGCTGC 5640
QY 5641 AAGGGAAACCTCGTGAAGAGTCCGGCACTGAAACACCCGCGGCCGCTCCACCTCCCGC 5700

Db 5641 AAGGGAAACCTCGTGAAGAGTCCGGCACTGAAACACCCGCGGCCGCTCCACCTCCCGC 5700
QY 5701 AGCAGCCCCAAAGCAGGAGGCCACCCACCGTACAAAGAGCACAATCACCAGCGCTGGCC 5760
Db 5701 AGCAGCCCCAAAGCAGGAGGCCACCCACCGTACAAAGAGCACAATCACCAGCGCTGGCC 5760
QY 5761 TCCAGCCAGCGCCGCCGAAGGCCCGCAGCCACCCGCGAGAGCAAGCAACCCCAACCCG 5820
Db 5761 TCCAGCCAGCGCCGCCGAAGGCCCGCAGCCACCCGCGAGAGCAAGCAACCCCAACCCG 5820
QY 5821 TACCGAGGGGGCGGACCGAGCTGCGCAGGCAAGTCTCTCGCCGCCGCCCTGGAGCGCA 5880
Db 5821 TACCGAGGGGGCGGACCGAGCTGCGCAGGCAAGTCTCTCGCCGCCGCCCTGGAGCGCA 5880
QY 5881 GAGAAGTCCCGCCCGCGATGCTCAAGCAAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5940
Db 5881 GAGAAGTCCCGCCCGCGATGCTCAAGCAAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5940
QY 5941 GAAGACAGCAGCAGGGGCCCGCTGCGGAGCGCTGCGGAGCCGCTGAGGACCCCGCTGCCAGGTG 6000
Db 5941 GAAGACAGCAGCAGGGGCCCGCTGCGGAGCGCTGCGGAGCCGCTGAGGACCCCGCTGCCAGGTG 6000
QY 6001 AACAAAGGGAAGAGCGCAGAGTGCCTCTCAAGTTTTCACGGTTTAACTGTCACTTATTAT 6060
Db 6001 AACAAAGGGAAGAGCGCAGAGTGCCTCTCAAGTTTTCACGGTTTAACTGTCACTTATTAT 6060
QY 6061 GACTGGAATTAAGCTGGACACACCTGCCAGCTTAACTGTCACTGTCACTGTCACTGTCACT 6120
Db 6061 GACTGGAATTAAGCTGGACACACCTGCCAGCTTAACTGTCACTGTCACTGTCACTGTCACT 6120
QY 6121 CTGAATCGAGAAATCCCGGACAGAGTGTGAAAGTCTGTTCTGAGAACAGATTATTCCTGA 6180
Db 6121 CTGAATCGAGAAATCCCGGACAGAGTGTGAAAGTCTGTTCTGAGAACAGATTATTCCTGA 6180
QY 6181 GCAGAGTTCATGTGACTTCTTAGACGTGTGACTTAAATAATGCTTAAAGCTTAAAGCTTAAAG 6240
Db 6181 GCAGAGTTCATGTGACTTCTTAGACGTGTGACTTAAATAATGCTTAAAGCTTAAAGCTTAAAG 6240
QY 6241 CAGCCACTCTGCTTACAAAAGAGTACTTGTAGTGCATGACTGTAGAAACAAATGTAA 6300
Db 6241 CAGCCACTCTGCTTACAAAAGAGTACTTGTAGTGCATGACTGTAGAAACAAATGTAA 6300
QY 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTTCCCTGGAGCC 6360
Db 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTTCCCTGGAGCC 6360
QY 6361 GAGAGCAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6420
Db 6361 GAGAGCAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6420
QY 6421 TAGCCAGGCTCAACAGGAGCAAGTGGCTGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6480
Db 6421 TAGCCAGGCTCAACAGGAGCAAGTGGCTGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6480
QY 6481 ATGTGTCCTCAGTCCCTGTTTCCAGAAATTTTCTGCGAAAGGAGTTAGCAATTCATTT 6540
Db 6481 ATGTGTCCTCAGTCCCTGTTTCCAGAAATTTTCTGCGAAAGGAGTTAGCAATTCATTT 6540
QY 6541 TTGGCTTTAAGAAAAATCGAGAATGTAGGTTTGA 6574
Db 6541 TTGGCTTTAAGAAAAATCGAGAATGTAGGTTTGA 6574

RESULT 2

US-10-325-430-10
; Sequence 10, Application US/10325430
; Publication No. US2003015325A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING


```

; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838,336 and 52908
; FILE REFERENCE: MP101-294PIRNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-325-430-10

```

Query Match	100.0%;	Score 6574;	DB 15;	Length 6574;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6574;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGAATGCT	60	
DB	1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGAATGCT	60	
QY	61	GGTGTGCTGAACCCATATGCCAGCGGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAA	120	
DB	61	GGTGTGCTGAACCCATATGCCAGCGGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAA	120	
QY	121	CCACCCCTTATGACTCAACAGCAGATGCTCCTCTTTCCGAGAAAGGATATTAGATGCC	180	
DB	121	CCACCCCTTATGACTCAACAGCAGATGCTCCTCTTTCCGAGAAAGGATATTAGATGCC	180	
QY	181	CTCTTTGTTCTTTTGAAGAAATCAGCTCAGCTGCTCTGTATGAAGATTAAGCAGTGAAC	240	
DB	181	CTCTTTGTTCTTTTGAAGAAATCAGCTCAGCTGCTCTGTATGAAGATTAAGCAGTGAAC	240	
QY	241	AACCTTTGTCGGGAAGTATCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCA	300	
DB	241	AACCTTTGTCGGGAAGTATCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCA	300	
QY	301	AAGGACTTCCGAAGTTCAGAACTCTTTGAGTTTGGTTCACCTTTGCTGAAGTGCAGGTGGTA	360	
DB	301	AAGGACTTCCGAAGTTCAGAACTCTTTGAGTTTGGTTCACCTTTGCTGAAGTGCAGGTGGTA	360	
QY	361	AGAGAGAAACCAACCGGGGACATCTATGCTATGAAAGTGTATGAAGAAGAAGGCTTTAATG	420	
DB	361	AGAGAGAAACCAACCGGGGACATCTATGCTATGAAAGTGTATGAAGAAGAAGGCTTTAATG	420	
QY	421	GCCAGAGAGCAGGTTTCATTTTTTGGAGAAAGCGGAACATATTATCTCGAAGCACAAAGC	480	
DB	421	GCCAGAGAGCAGGTTTCATTTTTTGGAGAAAGCGGAACATATTATCTCGAAGCACAAAGC	480	
QY	481	CCGTGGATCCCCCAATTACAGTATGCCCTTTTCAGGACAAAATCACCCTTTATCTGATCGAG	540	
DB	481	CCGTGGATCCCCCAATTACAGTATGCCCTTTTCAGGACAAAATCACCCTTTATCTGATCGAG	540	
QY	541	GAATATCAGCCTCGAGGGGACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600	
DB	541	GAATATCAGCCTCGAGGGGACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600	
QY	601	GAATACTGATACAGTTTATACCTAGCTGAGCTGATTTTGGCTGTTTCACAGCTTCACTCTG	660	
DB	601	GAATACTGATACAGTTTATACCTAGCTGAGCTGATTTTGGCTGTTTCACAGCTTCACTCTG	660	
QY	661	ATCGGATACCTGATCGAGACATCAAGCCTGAGAACATTCCTGTTGACCGCACAGGACAC	720	
DB	661	ATCGGATACCTGATCGAGACATCAAGCCTGAGAACATTCCTGTTGACCGCACAGGACAC	720	
QY	721	ATCAAGCTGCTGATTTTGGATCTGCGCGGAATGAATTCACAAACAGATGGTGAATGCC	780	
DB	721	ATCAAGCTGCTGATTTTGGATCTGCGCGGAATGAATTCACAAACAGATGGTGAATGCC	780	
QY	781	AAACTCCCGATTGGGACCCAGATTACATGGCTCCTCGAAGTGTGACTGTGATGAACCGG	840	

781	DB	AAACTCCGGAATGGGACCCAGATTA	CATCGCTCTCGAAGTGTGATGAACGGG	840
841	QY	GATGGAAAAGCACCTACGGCCCTGGACTGTGACTGGTGGT	CAGTGGCGGTGATTGCCTAT	900
841	DB	GATGGAAAAGCACCTACGGCCCTGGACTGTGACTGGTGGT	CAGTGGCGGTGATTGCCTAT	900
901	QY	GAGATGATTTATGGGAGATCCCTCTTCGACAGGGAACT	CTGCCAGAACTTCAATAAC	960
901	DB	GAGATGATTTATGGGAGATCCCTCTTCGACAGGGAACT	CTGCCAGAACTTCAATAAC	960
961	QY	ATTATGAATTTCCAGGGGTTTTGAAATTTCCAGATGAC	CCCCAAAGTGCAGTACTTT	1020
961	DB	ATTATGAATTTCCAGGGGTTTTGAAATTTCCAGATGAC	CCCCAAAGTGCAGTACTTT	1020
1021	QY	CTTGATCTGATCAAAAGCTTGTTGTGGCGCAGAAAGAG	ACTGAAGTTTGAAGTCTTT	1080
1021	DB	CTTGATCTGATCAAAAGCTTGTTGTGGCGCAGAAAGAG	ACTGAAGTTTGAAGTCTTT	1080
1081	QY	TGCTGCCATCTTTCTCTCTAAATTTGACTGGAAACA	ACTCGTAACCTCTCTCCCCC	1140
1081	DB	TGCTGCCATCTTTCTCTCTAAATTTGACTGGAAACA	ACTCGTAACCTCTCTCCCCC	1140
1141	QY	TTGCTTCCCAACCTCAAGTCTGACGATGACACCTC	CAATTTTGAATGACACAGAGAAT	1200
1141	DB	TTGCTTCCCAACCTCAAGTCTGACGATGACACCTC	CAATTTTGAATGACACAGAGAAT	1200
1201	QY	TCGTGGGTTTATCTCTCTCGTCCGACGTGAGCCCT	CAGGCTTCTCGGTTGAAGACTG	1260
1201	DB	TCGTGGGTTTATCTCTCTCGTCCGACGTGAGCCCT	CAGGCTTCTCGGTTGAAGACTG	1260
1261	QY	CCGTTTGTGGGTTTTCGTACAGCAAGGCACTGGGA	TCTCGTGAATCTGAGTCTGTT	1320
1261	DB	CCGTTTGTGGGTTTTCGTACAGCAAGGCACTGGGA	TCTCGTGAATCTGAGTCTGTT	1320
1321	QY	GTGTCGGTCTGACATCCCTGCGCAAGACTAGCT	CCATCGAAGAAATCTTCTCATCAA	1380
1321	DB	GTGTCGGTCTGACATCCCTGCGCAAGACTAGCT	CCATCGAAGAAATCTTCTCATCAA	1380
1381	QY	AGCAAAAGAGCTACAAGACTCTCAGGACAAGTGT	CACAAGATGGACGAGAAATGACCCGG	1440
1381	DB	AGCAAAAGAGCTACAAGACTCTCAGGACAAGTGT	CACAAGATGGACGAGAAATGACCCGG	1440
1441	QY	TTACATCGGAGAGTGTGAGGTTGGAGGCTGTGCT	TAGTCAGAAAGAGGTGGAGCTGAAG	1500
1441	DB	TTACATCGGAGAGTGTGAGGTTGGAGGCTGTGCT	TAGTCAGAAAGAGGTGGAGCTGAAG	1500
1501	QY	GCCTCTGAGACTCAGAGATCCCTCTTGGACGAG	ACTTCTACCTACATCACAAGATGC	1560
1501	DB	GCCTCTGAGACTCAGAGATCCCTCTTGGACGAG	ACTTCTACCTACATCACAAGATGC	1560
1561	QY	AGTAGCTTTAAAGCGAAGTTTGGAGCAAGCA	CGGATGGAGGTGTCCAGGAGGATGACAAA	1620
1561	DB	AGTAGCTTTAAAGCGAAGTTTGGAGCAAGCA	CGGATGGAGGTGTCCAGGAGGATGACAAA	1620
1621	QY	GCATCGAGCTTCTCCATGATATCAGAGACGAG	CGGAGCTCCAAAGAAATCAAGAG	1680
1621	DB	GCATCGAGCTTCTCCATGATATCAGAGACGAG	CGGAGCTCCAAAGAAATCAAGAG	1680
1681	QY	CAGGAGTACCAGGCTCAAGTGGAAAGAAATG	AGTTGATGAATCAGTTTGGAAAGGAT	1740
1681	DB	CAGGAGTACCAGGCTCAAGTGGAAAGAAATG	AGTTGATGAATCAGTTTGGAAAGGAT	1740
1741	QY	CTTGCTCAGACAGAGACGAGTGTATCTCTAG	GAATCTGAGCTGAGAGTCTCGGCTT	1800
1741	DB	CTTGCTCAGACAGAGACGAGTGTATCTCTAG	GAATCTGAGCTGAGAGTCTCGGCTT	1800
1801	QY	GCTGCTCAAGAAATTAAGCGGAAAGCGACA	GAATGTTCAGATAAATCTTTGAAGGCTAAG	1860
1801	DB	GCTGCTCAAGAAATTAAGCGGAAAGCGACA	GAATGTTCAGATAAATCTTTGAAGGCTAAG	1860
1861	QY	GATCAAGGGGAAGCCTGAAGTGGGAGAAATAT	GCAGAACTGGAGAAATCAATGCTGAGCAG	1920
1861	DB	GATCAAGGGGAAGCCTGAAGTGGGAGAAATAT	GCAGAACTGGAGAAATCAATGCTGAGCAG	1920

QY 1921 CAGCTCAAAATTCAGGAGCTCCAGAGAACTCGAGAGAGGCTCCAAAGGAGCGAGCCGAG 1980
DB 1921 CAGCTCAAAATTCAGGAGCTCCAGAGAACTCGAGAGAGGCTCCAAAGGAGCGAGCCGAG 1980
QY 1981 AGGAGCTCGAGAGAGCTCGAGAACCGAGAGGATCTCTTCTGAAGGCATCAGAAAGAGCTG 2040
DB 1981 AGGAGCTCGAGAGAGCTCGAGAACCGAGAGGATCTCTTCTGAAGGCATCAGAAAGAGCTG 2040
QY 2041 GTGGAAGCTGAGGAACGCCGCCATCTCTGAGAAACAAAGTAAAGAGACTAGAGACCATG 2100
DB 2041 GTGGAAGCTGAGGAACGCCGCCATCTCTGAGAAACAAAGTAAAGAGACTAGAGACCATG 2100
QY 2101 GAGCGTAGAGAAACAGACTGAGAGGTGACATCCAGACAAATCCCAACAGATCCAGAG 2160
DB 2101 GAGCGTAGAGAAACAGACTGAGAGGTGACATCCAGACAAATCCCAACAGATCCAGAG 2160
QY 2161 ATGGCTGATAAATCTCGAGCTCGAAGGAAACATCGGAGGCCCAAGTCTCAGCCAG 2220
DB 2161 ATGGCTGATAAATCTCGAGCTCGAAGGAAACATCGGAGGCCCAAGTCTCAGCCAG 2220
QY 2221 CACTAGAGTGCACCTGAACAGAAAGACGACACTATGAGGAAAGATTAAGTGTG 2280
DB 2221 CACTAGAGTGCACCTGAACAGAAAGACGACACTATGAGGAAAGATTAAGTGTG 2280
QY 2281 GACAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAACATGATGAGAGA 2340
DB 2281 GACAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAACATGATGAGAGA 2340
QY 2341 CAGCAGAGAGGCCATGAGAAAGGCGAAATTTCTCAGCGAACAGAGGCGATGATCAAT 2400
DB 2341 CAGCAGAGAGGCCATGAGAAAGGCGAAATTTCTCAGCGAACAGAGGCGATGATCAAT 2400
QY 2401 GCTATGATTCGAAGATCAGATCCCTGGACAGAGGATTTGGAACCTGCTGAAGCAAT 2460
DB 2401 GCTATGATTCGAAGATCAGATCCCTGGACAGAGGATTTGGAACCTGCTGAAGCAAT 2460
QY 2461 AAATCTGACAAATAGCAGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATG 2520
DB 2461 AAATCTGACAAATAGCAGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATG 2520
QY 2521 ATTCTGAACTCAGCAACAGAAATTTTACCTGGACAGACAGGCTCGGAGGTTGAGGCC 2580
DB 2521 ATTCTGAACTCAGCAACAGAAATTTTACCTGGACAGACAGGCTCGGAGGTTGAGGCC 2580
QY 2581 CAGAAACGAAAACTGGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACTGACCAAG 2640
DB 2581 CAGAAACGAAAACTGGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACTGACCAAG 2640
QY 2641 AATCGGCTGCTGAACTGAGACAGATTCGGGAGGTCAGTCTAGAGACAGAGAGCAG 2700
DB 2641 AATCGGCTGCTGAACTGAGACAGATTCGGGAGGTCAGTCTAGAGACAGAGAGCAG 2700
QY 2701 AAATCGAGCTCAAGCGCCAGCTCAGAGCTACAGCTCTCCCTGCGAGGCGCAGTCA 2760
DB 2701 AAATCGAGCTCAAGCGCCAGCTCAGAGCTACAGCTCTCCCTGCGAGGCGCAGTCA 2760
QY 2761 CAGTTGACAGCCCTGAGGCTGCA CGGCGGCCCTGGAGAGCCAGTTCGCCAGCGAAG 2820
DB 2761 CAGTTGACAGCCCTGAGGCTGCA CGGCGGCCCTGGAGAGCCAGTTCGCCAGCGAAG 2820
QY 2821 ACAGAGCTGGAAGAGACCAAGCAGAAAGCTGAAGAGGAGATCCAGGCACTCAGCGCACAT 2880
DB 2821 ACAGAGCTGGAAGAGACCAAGCAGAAAGCTGAAGAGGAGATCCAGGCACTCAGCGCACAT 2880
QY 2881 AGAGATGAAATCCAGCGCAAAATTTGATGCTCTTCGPAACAGCTGATCTGTAATCAAGAC 2940
DB 2881 AGAGATGAAATCCAGCGCAAAATTTGATGCTCTTCGTAACAGCTGATCTGTAATCAAGAC 2940
QY 2941 CTGGAGGAGCAGCTAAACAGCTGACCGAGGACAGCTGAACTCAACCAACCAACTTC 3000
DB 2941 CTGGAGGAGCAGCTAAACAGCTGACCGAGGAGCAAGCTGAACTCAACCAACCAACTTC 3000

QY 3001 TACTTGTCCAACAACTCGATGAGGCTTCTGGCCCAACGACGAGATTGTACAACTCGCA 3060
DB 3001 TACTTGTCCAACAACTCGATGAGGCTTCTGGCCCAACGACGAGATTGTACAACTCGCA 3060
QY 3061 AGTGAAGTGAACCATCTCCGCCGGAGATCAGGAAACGAGAGATGCACTTACAGCCAG 3120
DB 3061 AGTGAAGTGAACCATCTCCGCCGGAGATCAGGAAACGAGAGATGCACTTACAGCCAG 3120
QY 3121 AAGCAACAGATGAGAGGCTCTGAAGACCACTGACCAATGCTGAGGAAACAGGTCATGGAT 3180
DB 3121 AAGCAACAGATGAGAGGCTCTGAAGACCACTGACCAATGCTGAGGAAACAGGTCATGGAT 3180
QY 3181 TTGAGGCGCTTAAACGATGAGCTCTAGAAAAGAGCGGAGTGGGAGGCTCGAGAGC 3240
DB 3181 TTGAGGCGCTTAAACGATGAGCTCTAGAAAAGAGCGGAGTGGGAGGCTCGAGAGC 3240
QY 3241 GTCTGGGTGATGAGAAATCCAGTTTGAAGTGTGGGTTCGAGAGCTGCAAGAAATGCTG 3300
DB 3241 GTCTGGGTGATGAGAAATCCAGTTTGAAGTGTGGGTTCGAGAGCTGCAAGAAATGCTG 3300
QY 3301 GACACCGAGAAACAGAGCGGCGAGCCGATCAGCGGATCACCGAGTCTCCAGGCTG 3360
DB 3301 GACACCGAGAAACAGAGCGGCGAGCCGATCAGCGGATCACCGAGTCTCCAGGCTG 3360
QY 3361 GTGAGCTGCGCAGTGAAGGAGCAAGGCTGAGATTTCTGCTCTGCGAGAGGCTCTCAAA 3420
DB 3361 GTGAGCTGCGCAGTGAAGGAGCAAGGCTGAGATTTCTGCTCTGCGAGAGGCTCTCAAA 3420
QY 3421 GAGCAGAGCTGAAGCGCGAGAGGCTCTCTGACAAAGCTCAATGACCTGGAGAGAGCAT 3480
DB 3421 GAGCAGAGCTGAAGCGCGAGAGGCTCTCTGACAAAGCTCAATGACCTGGAGAGAGCAT 3480
QY 3481 GCTATGCTTGAATGAAATGCCAAGCTTACAGCAGAAAGCTGGAGACTGAACAGAGCTC 3540
DB 3481 GCTATGCTTGAATGAAATGCCAAGCTTACAGCAGAAAGCTGGAGACTGAACAGAGCTC 3540
QY 3541 AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTAACAGCAGAGATGACCTGCAAGAAAT 3600
DB 3541 AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTAACAGCAGAGATGACCTGCAAGAAAT 3600
QY 3601 CACATTTTCCGTCTGACTCAAGGACTGCAAGAGCTTAGATCGGCTGATCTACTGAAG 3660
DB 3601 CACATTTTCCGTCTGACTCAAGGACTGCAAGAGCTTAGATCGGCTGATCTACTGAAG 3660
QY 3661 ACAGAAAGAGTGA CTGGAGTATCAGCTGGAACAACTCAGGTTCTCTATTCTCATGAA 3720
DB 3661 ACAGAAAGAGTGA CTGGAGTATCAGCTGGAACAACTCAGGTTCTCTATTCTCATGAA 3720
QY 3721 AAGGTGAATGGAAGCCTATTTCTCAACAAACCAATCTCATTTCTGCAAGCC 3780
DB 3721 AAGGTGAATGGAAGCCTATTTCTCAACAAACCAATCTCATTTCTGCAAGCC 3780
QY 3781 AAAATGACCAACCTGCTAAAAAGAAAAAGGGTTATTTAGTCCGACGGAAAGAGGCCCT 3840
DB 3781 AAAATGACCAACCTGCTAAAAAGAAAAAGGGTTATTTAGTCCGACGGAAAGAGGCCCT 3840
QY 3841 GCTTTACCCACACAGGTTCTCTGCGAGTACAATGAGCTGAAGCTGGCCCTGGAGAGGAG 3900
DB 3841 GCTTTACCCACACAGGTTCTCTGCGAGTACAATGAGCTGAAGCTGGCCCTGGAGAGGAG 3900
QY 3901 AAAGCTCGCTGTGAGAGCTAGAGGAGCCCTTTCAGAAAGCCCGCATCGAGCTCCGCTCC 3960
DB 3901 AAAGCTCGCTGTGAGAGCTAGAGGAGCCCTTTCAGAAAGCCCGCATCGAGCTCCGCTCC 3960
QY 3961 GCCCGGAGGAGGCTGCCACCGCAAGCAACCGAACCCACCACTCCAGCCAGCC 4020
DB 3961 GCCCGGAGGAGGCTGCCACCGCAAGCAACCGAACCCACCACTCCAGCCAGCC 4020
QY 4021 ACCGCGAGGAGCAGATCGCCATGTCGGCCATGTCGGGTGCGCAGAGCACCAAGCCAGT 4080
DB 4021 ACCGCGAGGAGCAGATCGCCATGTCGGCCATGTCGGGTGCGCAGAGCACCAAGCCAGT 4080
QY 4081 GCCATGAGCCTGTGTGGCCCGCCATCCAGCGCAGAAAGAGGTTCTTCACTCCAGAGGAA 4140

QY 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360
Db 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360
QY 6361 GAGAGACAAATCTCTGTGTTGATTTTGAAGGACAGGCAAGACCAACACTGTATTAGTTCCA 6420
Db 6361 GAGAGACAAATCTCTGTGTTGATTTTGAAGGACAGGCAAGACCAACACTGTATTAGTTCCA 6420
QY 6421 TAGCCAGGCTTCAACAGGAGCAAGTGGCTGGCCCTTAAAACACACAGATGACTGGAATG 6480
Db 6421 TAGCCAGGCTTCAACAGGAGCAAGTGGCTGGCCCTTAAAACACACAGATGACTGGAATG 6480
QY 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGAGTTAGCAATTCATTT 6540
Db 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGAGTTAGCAATTCATTT 6540
QY 6541 TTGGCTTAAGAAAAATCGAGAAATGTAGTTTGA 6574
Db 6541 TTGGCTTAAGAAAAATCGAGAAATGTAGTTTGA 6574

RESULT 3

US-10-325-430-11
; Sequence 11, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: 32838,336 and 52908
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 6162
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6162)
US-10-325-430-11

Query Match 93.7%; Score 6162; DB 15; Length 6162;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 ATGTGGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGTGCTCTGAACCCATT 78
Db 1 ATGTGGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGTGCTCTGAACCCATT 60
QY 79 GCCAGCGGCTCCAGGCTGAATCTTTCTCCAGGGAAACCAACCTTTATGACTCAA 138
Db 61 GCCAGCGGCTCCAGGCTGAATCTTTCTCCAGGGAAACCAACCTTTATGACTCAA 120
QY 139 CAGCAGATGTCCTCTTTCCAGAGGGAATATTAGATGCCCTCTTTGCTCTTTGAA 198
Db 121 CAGCAGATGTCCTCTTTCCAGAGGGAATATTAGATGCCCTCTTTGCTCTTTGAA 180
QY 199 GAATGAGTCAGCTGCTGTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAGAT 258
Db 181 GAATGAGTCAGCTGCTGTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAGAT 240
QY 259 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCCGCAAGGACTTCGAAGTCAGA 318
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCCGCAAGGACTTCGAAGTCAGA 300

QY 319 AGTCTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGTTAAAGAGAGAAAGCAACCGG 378
Db 301 AGTCTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGTTAAAGAGAGAAAGCAACCGG 360
QY 379 GACATCTATGCTATGAAGTGTATGAAGAAGAGCTTTATTTGCCCCAGGAGAGGTTTCA 438
Db 361 GACATCTATGCTATGAAGTGTATGAAGAAGAGCTTTATTTGCCCCAGGAGAGGTTTCA 420
QY 439 TTTTTCAGAGAGAGCGGAACATATTATCTCGAAGCAACAAGCCGCTGGATCCCCCAATTA 498
Db 421 TTTTTCAGAGAGAGCGGAACATATTATCTCGAAGCAACAAGCCGCTGGATCCCCCAATTA 480
QY 499 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATGAGGAATATCAGCCTGGAGG 558
Db 481 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATGAGGAATATCAGCCTGGAGG 540
QY 559 GACTTGTGTCACCTTTTGAATAGATATGAGACAGTATGAAACCTGATACACATTT 618
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGACAGTATGAAACCTGATACACATTT 600
QY 619 TACCTAGCTAGCTGATTTTGGCTGTTTACAGCGCTTCATCTGATGGGATACGTGCATCGA 678
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTACAGCGCTTCATCTGATGGGATACGTGCATCGA 660
QY 679 GACATCAAGCTGAGAACATTCGTTGACGGACAGGACACATCAGCTGGTGGATTT 738
Db 661 GACATCAAGCTGAGAACATTCGTTGACGGACAGGACACATCAGCTGGTGGATTT 720
QY 739 GGATCTGCGCGAAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTTGGACC 798
Db 721 GGATCTGCGCGAAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTTGGACC 780
QY 799 CCAGATTACATGGCTCTCTGAAGTGTGACTGTGATGAACCGGGATGAAAAAGGACCTAC 858
Db 781 CCAGATTACATGGCTCTCTGAAGTGTGACTGTGATGAACCGGGATGAAAAAGGACCTAC 840
QY 859 GGCTCGACTGTGACTGTGCTGAGTGGGCTGATTGCCCTATCAGATGATTTATGGGAGA 918
Db 841 GGCTCGACTGTGACTGTGCTGAGTGGGCTGATTGCCCTATCAGATGATTTATGGGAGA 900
QY 919 TCCCCCTTCCGAGAGGAACTCTGCCAGAACCTTCAATAACATTTATGAATTTCCAGCGG 978
Db 901 TCCCCCTTCCGAGAGGAACTCTGCCAGAACCTTCAATAACATTTATGAATTTCCAGCGG 960
QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATTTCAAGC 1038
Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATTTCAAGC 1020
QY 1039 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTGTGAAGTCTTTGCTGCCACCTTTCTTC 1098
Db 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTGTGAAGTCTTTGCTGCCACCTTTCTTC 1080
QY 1099 TCTAAAATTGACTGGAAACAATTCGTAACCTCTCCCTCCCTCCCTCGTCCCAACCTCAAG 1158
Db 1081 TCTAAAATTGACTGGAAACAATTCGTAACCTCTCTCCCTCCCTCGTCCCAACCTCAAG 1140
QY 1159 TCTCAGCATGACACTCCAAATTTTGTATGAACAGAGAAATTCGTGGGTTTCATCTCT 1218
Db 1141 TCTCAGCATGACACTCCAAATTTTGTATGAACAGAGAAATTCGTGGGTTTCATCTCT 1200
QY 1219 CCCTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTTCG 1278
Db 1201 CCCTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTTCG 1260
QY 1279 TACAGAGGACACTGGGGATTTCTGGTAGATCTGAGTCTGTGTGCGGTCTGGACTCC 1338
Db 1261 TACAGAGGACACTGGGGATTTCTGGTAGATCTGAGTCTGTGTGCGGTCTGGACTCC 1320
QY 1339 CCTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATAAAAGCAAGAGCTACAAGAC 1398
Db 1321 CCTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATAAAAGCAAGAGCTACAAGAC 1380
QY 1399 TCTCAGGACAAAGTGTACAAGATGGAGCAGGAAATGACCCGGTTTACATCGGAGAGTCA 1458

Db 1381 TCTCAGGACAAAGTGTCACAAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGCA 1440
QY 1459 GAGGTGGAGCTGCTTGTAGTCAGAAAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518
Db 1441 GAGGTGGAGCTGCTTGTAGTCAGAAAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
QY 1519 TCCCTCCTGGAGCAGGACCTTCTACCTACATCAACAGAAATGCAAGGCTTAAAGGCAAGT 1578
Db 1501 TCCCTCCTGGAGCAGGACCTTCTACCTACATCAACAGAAATGCAAGGCTTAAAGGCAAGT 1560
QY 1579 TTGGAGCAAGCAGGATGAGTGTCCAGAGAGGATGACAAAGCACTGAGCTTCTCCAT 1638
Db 1561 TTGGAGCAAGCAGGATGAGTGTCCAGAGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
QY 1639 GATATCAGAGACGAGCGGAGGCTTCCAAAGAAATCAAAAGAGCAGAGTACCAGGCTCAA 1698
Db 1621 GATATCAGAGACGAGCGGAGGCTTCCAAAGAAATCAAAAGAGCAGAGTACCAGGCTCAA 1680
QY 1699 GTGGAAGAAATCAGGTTGATGATGAATCAGTTTGGAGGATCTTTGTCTCAGCAAGAAGA 1758
Db 1681 GTGGAAGAAATCAGGTTGATGATGAATCAGTTTGGAGGATCTTTGTCTCAGCAAGAAGA 1740
QY 1759 CGGAGTGATCTCTAGCAATCTGAGCTGAGAGTCTCGGCTGCTGCTGAAGATTCAG 1818
Db 1741 CGGAGTGATCTCTAGCAATCTGAGCTGAGAGTCTCGGCTGCTGCTGAAGATTCAG 1800
QY 1819 CGGAAAGGCACAGAATGTCAGCATAAATCTGTTGAAGGCTTAAGGATCAAGGGAAGCCTGAA 1878
Db 1801 CGGAAAGGCACAGAATGTCAGCATAAATCTGTTGAAGGCTTAAGGATCAAGGGAAGCCTGAA 1860
QY 1879 GTGGAGATATGCGAAATCGAGAGATCAATGCTGAGAGCAGCTCAAAATTCAGAG 1938
Db 1861 GTGGAGATATGCGAAATCGAGAGATCAATGCTGAGAGCAGCTCAAAATTCAGAG 1920
QY 1939 CTCGAAAGAACTCGAGAGGCTCAAGAGGAGCGAGAGGAGCTGGAGAGCTG 1998
Db 1921 CTCGAAAGAACTCGAGAGGCTCAAGAGGAGCGAGAGGAGCTGGAGAGCTG 1980
QY 1999 CAGAACCGAGAGGATTCCTTCTGAAGGATCAAGAAAGAGCTGGTGAAGCTGAGGAAGC 2058
Db 1981 CAGAACCGAGAGGATTCCTTCTGAAGGATCAAGAAAGAGCTGGTGAAGCTGAGGAAGC 2040
QY 2059 CGCCATTCCTGGAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2118
Db 2041 CGCCATTCCTGGAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2100
QY 2119 CTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2178
Db 2101 CTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2160
QY 2179 GAGCTCGAGAGAAACATCGGGAGGCCCAAGTCTCAGGCCAGCACCTAGAAAGTGCACCTG 2238
Db 2161 GAGCTCGAGAGAAACATCGGGAGGCCCAAGTCTCAGGCCAGCACCTAGAAAGTGCACCTG 2220
QY 2239 AAACAGAAAGAGCAGCAGCTATCAGGAAAGATTTAAAGTGTGGACAAATCAGATAAGAAA 2298
Db 2221 AAACAGAAAGAGCAGCAGCTATCAGGAAAGATTTAAAGTGTGGACAAATCAGATAAGAAA 2280
QY 2299 GACCTGGCTGACAGAGACACTGAGAACATGATGACAGACACAGAGGAGGCCCCAT 2358
Db 2281 GACCTGGCTGACAGAGACACTGAGAACATGATGACAGACACAGAGGAGGCCCCAT 2340
QY 2359 GAGAGGCGCAAAATCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCCAAGATC 2418
Db 2341 GAGAGGCGCAAAATCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCCAAGATC 2400
QY 2419 AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAACTTGCAGCAAAATAGC 2478
Db 2401 AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAACTTGCAGCAAAATAGC 2460
QY 2479 AGTCTTTTACCCAAAGAACTAGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2538

Db 2461 AGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520
QY 2539 CAGAAATTTTACCTGGAGACACAGGCTGGAAAGTTGGAGGCCAGAGACCGAAACTGGAG 2598
Db 2521 CAGAAATTTTACCTGGAGACACAGGCTGGAAAGTTGGAGGCCAGAGACCGAAACTGGAG 2580
QY 2599 GAGCAGCTGGAGAGAGATCAGCCCAAGAACACAGTGACAAAGAAATCGGCTGCTGGAACTG 2658
Db 2581 GAGCAGCTGGAGAGAGATCAGCCCAAGAACACAGTGACAAAGAAATCGGCTGCTGGAACTG 2640
QY 2659 GAGCAAGATTTGGGAGAGTCACTTAGACACAGAGAGCAGAGAACTGGAGCTCAAGCGC 2718
Db 2641 GAGCAAGATTTGGGAGAGTCACTTAGACACAGAGAGCAGAGAACTGGAGCTCAAGCGC 2700
QY 2719 CAGCTCACAGAGTACAGCTCTCCCTGCAGAGCGGAGTCAAGTTGACAGCCCTGCGAG 2778
Db 2701 CAGCTCACAGAGTACAGCTCTCCCTGCAGAGCGGAGTCAAGTTGACAGCCCTGCGAG 2760
QY 2779 GTTGCACGGCGGCTTGGAGAGCAGCTTTCGCGAGCGAAGACAGAGCTGGAGAGAGACC 2838
Db 2761 GCTGCACGGCGGCTTGGAGAGCAGCTTTCGCGAGCGAAGACAGAGCTGGAGAGAGACC 2820
QY 2839 ACAGCAGAACTGAAGAGGAGATCCAGGCACTCAGGCACTACGSCACATAGAGATGAATCCAGCGC 2898
Db 2821 ACAGCAGAACTGAAGAGGAGATCCAGGCACTCAGGCACTACGSCACATAGAGATGAATCCAGCGC 2880
QY 2899 AAATTTGATCTCTTCGTAAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2958
Db 2881 AAATTTGATCTCTTCGTAAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2940
QY 2959 CAGCTGACCGAGACAACTGAACTCAACAAACAAACTTCTACTTGTCCAAACAACTC 3018
Db 2941 CAGCTGACCGAGACAACTGAACTCAACAAACAAACTTCTACTTGTCCAAACAACTC 3000
QY 3019 GATGAGGCTTCTCGGCGCAACGAGAGATGTAACTGCGAAAGTGAAGTGGACCACTTC 3078
Db 3001 GATGAGGCTTCTCGGCGCAACGAGAGATGTAACTGCGAAAGTGAAGTGGACCACTTC 3060
QY 3079 CGCGGAGATCACGGAACGAGAGATGACAGCTTACAGCCAGAGCAACAGATGGAGGCT 3138
Db 3061 CGCGGAGATCACGGAACGAGAGATGACAGCTTACAGCCAGAGCAACAGATGGAGGCT 3120
QY 3139 CTGAAGACCACTGACCATGCTGGAGGAAACAGTCAATGGATTTGGAGGCCCTTAAACGAT 3198
Db 3121 CTGAAGACCACTGACCATGCTGGAGGAAACAGTCAATGGATTTGGAGGCCCTTAAACGAT 3180
QY 3199 GAGCTGCTAGAAAAAGAGCGGAGTGGAGGCTTGGAGGAGCGCTCTGGGTGATGAGAAA 3258
Db 3181 GAGCTGCTAGAAAAAGAGCGGAGTGGAGGCTTGGAGGAGCGCTCTGGGTGATGAGAAA 3240
QY 3259 TCCGAGTTTGAAGTGTGGGTTGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGC 3318
Db 3241 TCCGAGTTTGAAGTGTGGGTTGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGC 3300
QY 3319 AGGCGAGAGCCGATCAGCGGATCAGCGAGTCTCGCGAGTGTGGAGCTGGCAGTGAAG 3378
Db 3301 AGGCGAGAGCCGATCAGCGGATCAGCGAGTCTCGCGAGTGTGGAGCTGGCAGTGAAG 3360
QY 3379 GAGCACAAGGCTGAGATTTCTGCTGCGAGAGCTCTCAAAGAGCAGAGCTGAAAGGCC 3438
Db 3361 GAGCACAAGGCTGAGATTTCTGCTGCGAGAGCTCTCAAAGAGCAGAGCTGAAAGGCC 3420
QY 3439 GAGAGCTCTCTCAAGGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3498
Db 3421 GAGAGCTCTCTGACAAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAAT 3480
QY 3499 GCCCGAAGCTTACAGCAGAGCTGGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAA 3558
Db 3481 GCCCGAAGCTTACAGCAGAGCTGGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAA 3540
QY 3559 GAGCAAGCCAAATTAAGCAGAGAGTGGAGCTGCAAGAAATCAACATTTTCCGCTGACT 3618
Db 3541 GAGCAAGCCAAATTAAGCAGAGAGTGGAGCTGCAAGAAATCAACATTTTCCGCTGACT 3600

3619 CAAGGACTGCAAGAACTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTTG 3678
|||||
3601 CAAGGACTGCAAGAACTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTTG 3660
|||||
3679 GAGTATCAGGTGAAACAACTTCAAGTTCTTCTATCTCATGAAAGGTGAAATGGAAGGC 3738
|||||
3661 GAGTATCAGGTGAAACAACTTCAAGTTCTTCTATCTCATGAAAGGTGAAATGGAAGGC 3720
|||||
3739 ACTATTCTCAACAAACCAACTCAATGATTTTCTGCAAGCCAAATGGAACCACTGCT 3798
|||||
3721 ACTATTCTCAACAAACCAACTCAATGATTTTCTGCAAGCCAAATGGAACCACTGCT 3780
|||||
3799 AAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAGGACCTTGCTTTACCCACACAGGTT 3858
|||||
3781 AAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAGGACCTTGCTTTACCCACACAGGTT 3840
|||||
3859 CCTCTCAGTACATCAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAG 3918
|||||
3841 CCTCTCAGTACATCAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAG 3900
|||||
3919 CTAGAGAAAGCCCTTCAGAAAGACCCGCATCGAGCTCCGGTCCGCCGGGAGGAAGCTGCC 3978
|||||
3901 CTAGAGAAAGCCCTTCAGAAAGACCCGCATCGAGCTCCGGTCCGCCGGGAGGAAGCTGCC 3960
|||||
3979 CACCGAAGACGACACCAACCCACACCCATCCAGCCGACCCACCGCGAGGACGAGATC 4038
|||||
3961 CACCGAAGACGACACCAACCCACACCCATCCAGCCGACCCACCGCGAGGACGAGATC 4020
|||||
4039 GCCATGTCGCCCATCTGCGGTCCGACAGACCCAGCCAGCTGCCATGAGCCCTGCTGGCC 4098
|||||
4021 GCCATGTCGCCCATCTGCGGTCCGACAGACCCAGCCAGCTGCCATGAGCCCTGCTGGCC 4080
|||||
4099 CCGCCATCCAGCCGCGAAGAGAGTCTTCAACTCAGAGGAAATTTAGTCCGGCTCTTAAG 4158
|||||
4081 CCGCCATCCAGCCGCGAAGAGAGTCTTCAACTCAGAGGAAATTTAGTCCGGCTCTTAAG 4140
|||||
4159 GAACGATGACACCAATATCTCTACCGATTCACAGTAGGACTGAAATGGAGCCACA 4218
|||||
4141 GAACGATGACACCAATATCTCTACCGATTCACAGTAGGACTGAAATGGAGCCACA 4200
|||||
4219 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGAACGCCAGGCATCCAAATGTCTCGAA 4278
|||||
4201 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGAACGCCAGGCATCCAAATGTCTCGAA 4260
|||||
4279 TGTTCAGGTGATGTGTACCCCAAGTGTCCAGTGTCTCCAGCCACCTGCGGCTTGCT 4338
|||||
4261 TGTTCAGGTGATGTGTACCCCAAGTGTCCAGTGTCTCCAGCCACCTGCGGCTTGCT 4320
|||||
4339 GCTGAATATGCCACACACTTCCAGAGGCTTCTGCGGTGACAAAATGAACTCCCAAGT 4398
|||||
4321 GCTGAATATGCCACACACTTCCAGAGGCTTCTGCGGTGACAAAATGAACTCCCAAGT 4380
|||||
4399 CTCAGACCAAGAGGCCACAGCAGCTTGGACCTGGAGAGGTGATGAGTGGCCAGG 4458
|||||
4381 CTCAGACCAAGAGGCCACAGCAGCTTGGACCTGGAGAGGTGATGAGTGGCCAGG 4440
|||||
4459 AATAACAAACGAGCAGCAAGGCTGGGACAGGATGATATGTCCTGGAGGGATCAAAA 4518
|||||
4441 AATAACAAACGAGCAGCAAGGCTGGGACAGGATGATATGTCCTGGAGGGATCAAAA 4500
|||||
4519 GTCTCATTTATGACAAATGAAGCAGAGAGTGGACAGAGCCCGGTGGAAGAAATTTGAG 4578
|||||
4501 GTCTCATTTATGACAAATGAAGCAGAGAGTGGACAGAGCCCGGTGGAAGAAATTTGAG 4560
|||||
4579 CTGTGCTTCCGACCGGGATGTATCTATTTCATGTTGGCTTGGTGTTCCTCGAATTCGCA 4638
|||||
4561 CTGTGCTTCCGACCGGGATGTATCTATTTCATGTTGGCTTGGTGTTCCTCGAATTCGCA 4620
|||||
4639 AATAACGCAAAAGCAAAAAAGCAGAGTGTGATGCTTAAATCTGCTTGGAAATCTCCCTGCTG 4698
|||||
4621 AATAACGCAAAAGCAAAAAAGCAGAGTGTGATGCTTAAATCTGCTTGGAAATCTCCCTGCTG 4680
|||||

4699 AAATCGAAAGGTGATGACCGCTAGACATGAATGCAACGCTGCCCTTCTAGTGACCAAGTG 4758
|||||
4681 AAATCGAAAGGTGATGACCGCTAGACATGAATGCAACGCTGCCCTTCTAGTGACCAAGTG 4740
|||||
4759 GTGTTGTGGGCAACCGAGGAGGCTCTACGCCCTGAATGCTCTTGAATAACTCCCTTAACC 4818
|||||
4741 GTGTTGTGGGCAACCGAGGAGGCTCTACGCCCTGAATGCTCTTGAATAACTCCCTTAACC 4800
|||||
4819 CATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATTAAGGACCTTGGGAAGCTA 4878
|||||
4801 CATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATTAAGGACCTTGGGAAGCTA 4860
|||||
4879 CTCAATGATAGCAGGAGAGCGGCACTGTGCTTGTGGACGTGAAGAAAGTGAACAG 4938
|||||
4861 CTCAATGATAGCAGGAGAGCGGCACTGTGCTTGTGGACGTGAAGAAAGTGAACAG 4920
|||||
4939 TCCCTGGCCCAAGTCCCACTGCTGCCAGCCGACATCTCACCCCAACATTTTTGAAGCT 4998
|||||
4921 TCCCTGGCCCAAGTCCCACTGCTGCCAGCCGACATCTCACCCCAACATTTTTGAAGCT 4980
|||||
4999 GTCAAGGGCTGCCACTTTGTTTGGGCGAGGCAAGATTGAGAACGGGCTCTGCATCTGTGA 5058
|||||
4981 GTCAAGGGCTGCCACTTTGTTTGGGCGAGGCAAGATTGAGAACGGGCTCTGCATCTGTGA 5040
|||||
5059 GCCATGCCAGCAAGTCTGCTATCTCCGCTACAACGAAACCTCAGCAAAATACTGCATC 5118
|||||
5041 GCCATGCCAGCAAGTCTGCTATCTCCGCTACAACGAAACCTCAGCAAAATACTGCATC 5100
|||||
5119 CGGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCATTCACCAATTCAGTATC 5178
|||||
5101 CGGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCATTCACCAATTCAGTATC 5160
|||||
5179 CTCAATCGAAACCAATTAATTTCTAGAAATCGCATGAAGCAGTACACGCTCGAGGAATTC 5238
|||||
5161 CTCAATCGAAACCAATTAATTTCTAGAAATCGCATGAAGCAGTACACGCTCGAGGAATTC 5220
|||||
5239 CTGATTAAGAAATGACCAATTCCTTTGGCACTGTGTGTTTGGCCCTCTTCCAAACAGCTTC 5298
|||||
5221 CTGATTAAGAAATGACCAATTCCTTTGGCACTGTGTGTTTGGCCCTCTTCCAAACAGCTTC 5280
|||||
5299 CTTGTCTCAATCGTGCAGGTGAACAGCGGAGGCGAGGAGAGGACTTGTGCTGTTTC 5358
|||||
5281 CTTGTCTCAATCGTGCAGGTGAACAGCGGAGGCGAGGAGAGGACTTGTGCTGTTTC 5340
|||||
5359 CACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGAGCTAGCCGACAGACGATCTCAAG 5418
|||||
5341 CACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGAGCTAGCCGACAGACGATCTCAAG 5400
|||||
5419 TGGAGTCCGTTACTTTTGGCTTTGCTTACAGAAACCTTATCTGTTTGTGACCACTTC 5478
|||||
5401 TGGAGTCCGTTACTTTTGGCTTTGCTTACAGAAACCTTATCTGTTTGTGACCACTTC 5460
|||||
5479 AACTCAGTTCGAAGTAAATTTAGATCCAGGCACGCTCTCAGCAGGACCCCTGCCGAGCG 5538
|||||
5461 AACTCAGTTCGAAGTAAATTTAGATCCAGGCACGCTCTCAGCAGGACCCCTGCCGAGCG 5520
|||||
5539 TAQCTGGACATCCGAAACCCGCGCTACTTGGGCGCTGCCATTTCTCAGAGGAGGATTTAC 5598
|||||
5521 TACCTGGACATCCGAAACCCGCGCTACTTGGGCGCTGCCATTTCTCAGAGGAGGATTTAC 5580
|||||
5599 TTGCGCTCTCATACAGGATTAATTAAGGTCTATTGCTGCAAGGAAACCTCTGTAAG 5658
|||||
5581 TTGCGCTCTCATACAGGATTAATTAAGGTCTATTGCTGCAAGGAAACCTCTGTAAG 5640
|||||
5659 GAGTTCGGCACTGAACACCAACCCGCGCTCCACTCTCCGCGAGCAGCCCAACAGCGA 5718
|||||
5641 GAGTTCGGCACTGAACACCAACCCGCGCTCCACTCTCCGCGAGCAGCCCAACAGCGA 5700
|||||
5719 GGCCCAACCGGATACACGAGCAGATCAACCAAGGGTGGCTCCAGCCAGCGCCGCC 5778
|||||
5701 GGCCCAACCGGATACACGAGCAGATCAACCAAGGGTGGCTCCAGCCAGCGCCGCC 5760
|||||
5779 GAGGCCCCCAGCCACCCGCGAGAGCCAAAGCACACCCACCGCTTACCCGAGGGCGGACC 5838
|||||

Db 5761 GAAGCCCCAGCCACCCGCGAGAGCAAGCAACCCACCGCTACCGCGAGGGCGGACC 5820
Qy 5839 GAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCTCTGGAGCGAGAGAAAGTCCCGCGCGG 5898
Db 5821 GAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCTCTGGAGCGAGAGAAAGTCCCGCGCGG 5880
Qy 5899 ATGCTCAGCAGCGGAGAGAGCGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGAGGGC 5958
Db 5881 ATGCTCAGCACGCGAGAGAGGGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGAGGGC 5940
Qy 5959 CGGCTGCTGCGGAGCGCTGAGGAGCCCGCTGCTGCCAGGTGAACAAGGGAAGAGGGCAG 6018
Db 5941 CGGCTGCTGCGGAGCGCTGAGGAGCCCGCTGCTGCCAGGTGAACAAGGGAAGAGGGCAG 6000
Qy 6019 AGTGCCCTCTCAAGTTTTCACGGTTTAAACACTGTCACCTATTATGACTGGAAATAAAAGCTG 6078
Db 6001 AGTGCCCTCTCAAGTTTTCACGGTTTAAACACTGTCACCTATTATGACTGGAAATAAAAGCTG 6050
Qy 6079 GACAACTGCCAGCTAACTGGTCACTGCTGAGGATCATCCAGCTGAATGGAGAAATCCGG 6138
Db 6061 GACAACTGCCAGCTAACTGGTCACTGCTGAGGATCATCCAGCTGAATGGAGAAATCCGG 6120
Qy 6139 CAGCAGGTTGAAAGTCTGTTCTGAGACAGATTTTCTGTA 6180
Db 6121 CAGCAGGTTGAAAGTCTGTTCTGAGAACAGATTTTCTGTA 6162

RESULT 4
US-10-017-216-3
; Sequence, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMAN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1e1 Human Myotonic Dystrophy Type Prot
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-216-3

Qy 319 AGTCTGTAGTGTGTGCTCACTTTGCTGAAGTCAGGTGGTAAAGAGAGAAAGCAACCGGG 378
Db 301 AGTCTGTAGTGTGTGCTCACTTTGCTGAAGTCAGGTGGTAAAGAGAGAAAGCAACCGGG 360
Qy 379 GACATCTATGCTATGAAAAGTGAAGAAAGAGGCTTTATTTGGCCCCAGAGCAGGTTTCA 438
Db 361 GACATCTATGCTATGAAAAGTGAAGAAAGAGGCTTTATTTGGCCCCAGAGCAGGTTTCA 420
Qy 439 TTTTGTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 498
Db 421 TTTTGTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480
Qy 499 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGAGGAATATCAGCTGAGGG 558
Db 481 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGAGGAATATCAGCTGAGGG 540
Qy 559 GACTTCTGTCACCTTTTGAATAGATAGAGGACCACTTAGATGAAGAAACCTGTATACAGTTT 618
Db 541 GACTTCTGTCACCTTTTGAATAGATAGAGGACCACTTAGATGAAGAAACCTGTATACAGTTT 600
Qy 619 TACTAGCTGAGCTGATTTTGGCTGTTTCAGCGGTTTCATCTGATGGGATACGTGCATCGA 678
Db 601 TACTAGCTGAGCTGATTTTGGCTGTTTCAGCGGTTTCATCTGATGGGATACGTGCATCGA 660
Qy 679 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAAGGACACATCAAGCTGGTGGATTTT 738
Db 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAAGGACACATCAAGCTGGTGGATTTT 720
Qy 739 GGATCTGCCCGGAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 798
Db 721 GGATCTGCCCGGAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
Qy 799 CCAGATTACATGCTCTCTGAAGTCTGACTGTGATGAACGGGGATGGAAGGACCTTAC 858
Db 781 CCAGATTACATGCTCTCTGAAGTCTGACTGTGATGAACGGGGATGGAAGGACCTTAC 840
Qy 859 GSCCTGGAGCTGACTGGTGGTGCAGTGGCGTGAATTCCTATGAGATGATTTATGGAGA 918
Db 841 GSCCTGGAGCTGACTGGTGGTGCAGTGGCGTGAATTCCTATGAGATGATTTATGGAGA 900
Qy 919 TCCCTCTTCGACAGAGGGAACCTCTGCCAGAACCTTCAATAACAATTTATGAATTTCCAGCGG 978
Db 901 TCCCTCTTCGACAGAGGGAACCTCTGCCAGAACCTTCAATAACAATTTATGAATTTCCAGCGG 960
Qy 979 TTTTGTGAATTTCCAGATGACCCCAAGTAGAGCAGTGACTTTCTTGTATCTGATTCAGAGC 1038
Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTAGAGCAGTGACTTTCTTGTATCTGATTCAGAGC 1020
Qy 1039 TTGTTGTGGCGGCAGAAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC 1098
Db 1021 TTGTTGTGGCGGCAGAAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC 1080
Qy 1099 TCTAAAATTGACTGGAAACAACATTCGTAACCTCTCCTCCCGCTTCGTTCCACCCCTCAAG 1158
Db 1081 TCTAAAATTGACTGGAAACAACATTCGTAACCTCTCCTCCCGCTTCGTTCCACCCCTCAAG 1140
Qy 1159 TCTGAGCAGTACACCTCCAAATTTTGAATGAACACAGAGAAGAAATTCGTTGGGTTTTCATCCTCT 1218
Db 1141 TCTGAGCAGTACACCTCCAAATTTTGAATGAACACAGAGAAGAAATTCGTTGGGTTTTCATCCTCT 1200
Qy 1219 CCGTGGCAGCTGAGCCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1278
Db 1201 CCGTGGCAGCTGAGCCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
Qy 1279 TAGCAGCAGCAGCTGGGATCTTGTAGATCTGAGTCTGTTGTGTCGGGTCTGGACTCC 1338
Db 1261 TAGCAGCAGCAGCTGGGATCTTGTAGATCTGAGTCTGTTGTGTCGGGTCTGGACTCC 1320
Qy 1339 CTTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1398
Db 1321 CTTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1380
Qy 1399 TCTCAGGACAAGTGTCTCAAGATGGAGGAGAAATGACCCGTTTACATCGGAGAGTGTCA 1458

1381 TCTCAGGCAAGTGTACACAGATGGAGCAGGAATGACCGGTATACATCGGAGAGTGCA 1440
1459 GAGGTGGAGGCTGTGTTAGTCAGAAAGGAGTGGAGCTGAAGCCCTCTCAGACTCAGAGA 1518
1441 GAGGTGGAGGCTGTGTTAGTCAGAAAGGAGTGGAGCTGAAGCCCTCTCAGACTCAGAGA 1500
1519 TCCTCTCTGGAGCAGGACCTTGCTACTACATCACAGAATGAGTAGTCTTAAAGGGAAGT 1578
1501 TCCTCTCTGGAGCAGGACCTTGCTACTACATCACAGAATGAGTAGTCTTAAAGGGAAGT 1560
1579 TTGGAGCAAGCAGGATGGAGGTGCCAGAGGAGTGAAGAGGATGACAAAGCACTGCTCCAT 1638
1561 TTGGAGCAAGCAGGATGGAGGTGCCAGAGGAGTGAAGAGGATGACAAAGCACTGCTCCAT 1620
1639 GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGCAGGAGTACAGGCTCAA 1698
1621 GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGCAGGAGTACAGGCTCAA 1680
1699 GTGGAAGAAATGAGTGTGATGATGATCAATCAATGTTGGAAGAGGATCTTGTCTCAGCAAGA 1758
1681 GTGGAAGAAATGAGTGTGATGATGATGATCAATCAATGTTGGAAGAGGATCTTGTCTCAGCAAGA 1740
1759 CGGAGTGTATCTTACCAATCTCAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1818
1741 CGGAGTGTATCTTACCAATCTCAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1800
1819 CGGAAGCAGCAAGATGTCAGCAATAAATCTTTGAAGGCTAAGGATCAAGGAGGCTGAA 1878
1801 CGGAAGCAGCAAGATGTCAGCAATAAATCTTTGAAGGCTAAGGATCAAGGAGGCTGAA 1860
1879 GTGGGAGATATCCGAAATCTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938
1861 GTGGGAGATATCCGAAATCTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
1939 CTCGAAGAGAACTGGAGAGGCTGCAAAAGGAGCGAGCCGAGAGGAGCTGAGAGGCTG 1998
1921 CTCGAAGAGAACTGGAGAGGCTGCAAAAGGAGCGAGCCGAGAGGAGCTGAGAGGCTG 1980
1999 CAGAACCGAGAGATCTTCTGAGGAGCATCAGAAAGAGCTGAGAGAGCTGAGAGAGCG 2058
1981 CAGAACCGAGAGATCTTCTGAGGAGCATCAGAAAGAGCTGAGAGAGCTGAGAGAGCG 2040
2059 CGCATTCTCTGGAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2118
2041 CGCATTCTCTGGAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2100
2119 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATAAATTCG 2178
2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATAAATTCG 2160
2179 GAGCTCGAAGAGAAAATCGGAGGCCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTG 2238
2161 GAGCTCGAAGAGAAAATCGGAGGCCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTG 2220
2239 AAACAGAAAGAGCAGCCTATGAGGAAAGATTAAGTGTGGACAATCAGATAAGAAA 2298
2221 AAACAGAAAGAGCAGCCTATGAGGAAAGATTAAGTGTGGACAATCAGATAAGAAA 2280
2299 GAGCTGGCTGACAAGAGACACTGGAGAACATGATGTCAGAGACACGAGAGGAGGCCAT 2358
2281 GAGCTGGCTGACAAGAGACACTGGAGAACATGATGTCAGAGACACGAGAGGAGGCCAT 2340
2359 GAGAGGGCAAAATCTCAGCAACAGAGAGGCGATGATCAATGCTATGATTCAGATC 2418
2341 GAGAGGGCAAAATCTCAGCAACAGAGAGGCGATGATCAATGCTATGATTCAGATC 2400
2419 AGATCCCTGGAAACAGAGGATTTGGAACTGCTGAGGCCAATAAATTCGACGAAATAGC 2478
2401 AGATCCCTGGAAACAGAGGATTTGGAACTGCTGAGGCCAATAAATTCGACGAAATAGC 2460
2479 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAA 2538

2461 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520
2539 CAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCCGAAACTGGAG 2598
2521 CAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCCGAAACTGGAG 2580
2599 GAGCAGCTGGAGAAAGATCAGCCACCAAGACACAGTGCACAAGAAATCGGCTGCTGGAAGCTG 2658
2581 GAGCAGCTGGAGAAAGATCAGCCACCAAGACACAGTGCACAAGAAATCGGCTGCTGGAAGCTG 2640
2659 GAGCAAGATTTGCGGAGGTCAGTCTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAGCGC 2718
2641 GAGCAAGATTTGCGGAGGTCAGTCTAGAGCAGCAGGAGCAGAACTGGAGCTCAAGCGC 2700
2719 CAGCTCAGAGCTACAGCTCTCCCTCAGAGGCGCGAGTCAAGTTCAGAGCCCTGCAG 2778
2701 CAGCTCAGAGCTACAGCTCTCCCTCAGAGGCGCGAGTCAAGTTCAGAGCCCTGCAG 2760
2779 GCTCAGCGCGCGCTGGAGAGCGAGCTTCGCCAGCGCAAGACAGAGCTGGAGAGACC 2838
2761 GCTCAGCGCGCGCTGGAGAGCGAGCTTCGCCAGCGCAAGACAGAGCTGGAGAGACC 2820
2839 ACAGCAGAACTGAAGAGGAGATCCAGGCACTCAGCGCAATAGAGATGAAATCCAGCGC 2898
2821 ACAGCAGAACTGAAGAGGAGATCCAGGCACTCAGCGCAATAGAGATGAAATCCAGCGC 2880
2899 AATTTTATGCTCTTCTGTAACAGCTGCTACTGTAATCAGACACCTGGAGGAGCAGTAAAC 2958
2881 AATTTTATGCTCTTCTGTAACAGCTGCTACTGTAATCAGACACCTGGAGGAGCAGTAAAC 2940
2959 CAGCTGACCGAGGACACGCTGAACTCAAACAACAACTTCTACTTTGCCAAACAATC 3018
2941 CAGCTGACCGAGGACACGCTGAACTCAAACAACAACTTCTACTTTGCCAAACAATC 3000
3019 GATGAGCTTTCTGGGCCAACGAGAGATGTTGTAACAATCGGAAGTGAAGTGGACCATCTC 3078
3001 GATGAGCTTTCTGGGCCAACGAGAGATGTTGTAACAATCGGAAGTGAAGTGGACCATCTC 3060
3079 CGCCGGAGATCAGGAAACGAGAGATGCACTTACCAGCCAGAACCAACGATGGAGGCT 3138
3061 CGCCGGAGATCAGGAAACGAGAGATGCACTTACCAGCCAGAACCAACGATGGAGGCT 3120
3139 CTGAAGACCACTGCTCAGCCATGCTGGAGAAACAGGTATGGAATTTGGAGGCCCTAAACGAT 3198
3121 CTGAAGACCACTGCTCAGCCATGCTGGAGAAACAGGTATGGAATTTGGAGGCCCTAAACGAT 3180
3199 GAGCTGCTAGAAAAGAGCGGAGTGGAGGCTGGAGGAGCGCTCTGGTGTATGAGAAA 3258
3181 GAGCTGCTAGAAAAGAGCGGAGTGGAGGCTGGAGGAGCGCTCTGGTGTATGAGAAA 3240
3259 TCCAGTTTGAAGTGTCCGTTCCAGAGCTGCAAGAAATGCTGGACACCGAGAAACAGAGC 3318
3241 TCCAGTTTGAAGTGTCCGTTCCAGAGCTGCAAGAAATGCTGGACACCGAGAAACAGAGC 3300
3319 AGGCGAGAGCCGATCAGCGGATCAGCGAGTCTCGCAGGTGGAGCTGGAGGAGTGAAG 3378
3301 AGGCGAGAGCCGATCAGCGGATCAGCGAGTCTCGCAGGTGGAGCTGGAGGAGTGAAG 3360
3379 GAGCAAGGCTGAGATTTCTCGCTGTCAGCAGGCTCTCAAGAGCAGAGCTGAAGGCC 3438
3361 GAGCAAGGCTGAGATTTCTCGCTGTCAGCAGGCTCTCAAGAGCAGAGCTGAAGGCC 3420
3439 GAGAGCTCTCTGACAGCTCAATGATCCCTGGAGAAAGAGCTGCTTGTGAAATGAAT 3498
3421 GAGAGCTCTCTGACAGCTCAATGATCCCTGGAGAAAGAGCTGCTTGTGAAATGAAT 3480
3499 GCCCAAGCTTACGACAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3558
3481 GCCCAAGCTTACGACAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540
3559 GAGCAGGCCAAATTAACGACGAGAGTGGAGCTGTCAGAAAAATCAGATTTTCCGCTGACT 3618
3541 GAGCAGGCCAAATTAACGACGAGAGTGGAGCTGTCAGAAAAATCAGATTTTCCGCTGACT 3600

QY 3619 CAAGGACTGCAAGAGCTCTAGATCGGCGTGATCTACTGAAGACAGAAAGAGTACTTG 3678
Db 3601 CAAGGACTGCAAGAGCTCTAGATCGGCGTGATCTACTGAAGACAGAAAGAGTACTTG 3660
QY 3679 GAGTATCAGCTGGAAAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGGC 3738
Db 3661 GAGTATCAGCTGGAAAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGGC 3720
QY 3739 ACTATTCTCAACACCAACCAATCTATTGATTTCTGCAAGCCCAAAATGGACCACTGCT 3798
Db 3721 ACTATTCTCAACACCAACCAATCTATTGATTTCTGCAAGCCCAAAATGGACCACTGCT 3780
QY 3799 AAAAAAAGAGGTTTATTTAGTCGACCGAAAGAGGACCTGCTTTACCCACACAGGTT 3858
Db 3781 AAAAAAAGAGGTTTATTTAGTCGACCGAAAGAGGACCTGCTTTACCCACACAGGTT 3840
QY 3859 CCTCTGCAGTACAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAG 3918
Db 3841 CCTCTGCAGTACAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAG 3900
QY 3919 CTAGAGGAAGCCCTTCAGAAAGCCCGCATCGAGCTCCGCTCCGCGGAGGAAAGCTGCC 3978
Db 3901 CTAGAGGAAGCCCTTCAGAAAGCCCGCATCGAGCTCCGCTCCGCGGAGGAAAGCTGCC 3960
QY 3979 CACCGCAAGCAACGGACCAACCAACCAATCCACGCCAGCCACCGCGAGGCAGAGATC 4038
Db 3961 CACCGCAAGCAACGGACCAACCAACCAATCCACGCCAGCCACCGCGAGGCAGAGATC 4020
QY 4039 GCCATGTCGCCCATCGTCGGTGCAGAGCACCGCCAGCCAGCGCATGAGCTGTGCC 4098
Db 4021 GCCATGTCGCCCATCGTCGGTGCAGAGCACCGCCAGCCAGCGCATGAGCTGTGCC 4080
QY 4099 CGGCCATCCAGCGCGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAG 4158
Db 4081 CGGCCATCCAGCGCGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAG 4140
QY 4159 GAACGATCCACCAACATATTCTCACCGATTCACGTTAGGACTGAACATGGAGCCACA 4218
Db 4141 GAACGATCCACCAACATATTCTCACCGATTCACGTTAGGACTGAACATGGAGCCACA 4200
QY 4219 AAGTGTGCTGTGTCGATACCGGTGCACTTTGACGCGCAGGCATCCAAATGTCTCGAA 4278
Db 4201 AAGTGTGCTGTGTCGATACCGGTGCACTTTGACGCGCAGGCATCCAAATGTCTCGAA 4260
QY 4279 TGTGAGTGAATGTGTCAACCCCAAGTGTCCACGTGCTTGCCAGCCACCTGCGGCTTGCT 4338
Db 4261 TGTGAGTGAATGTGTCAACCCCAAGTGTCCACGTGCTTGCCAGCCACCTGCGGCTTGCT 4320
QY 4339 GCTGAATATGCCACACACTTCACCGAGGCTTCTCCGTGACAAATGAATCCCCAGGT 4398
Db 4321 GCTGAATATGCCACACACTTCACCGAGGCTTCTCCGTGACAAATGAATCCCCAGGT 4380
QY 4399 CTCAGACCAAGGAGCCACGACGACTTGCACCTGGAGGGTGGATGAAGGTGCCCAGG 4458
Db 4381 CTCAGACCAAGGAGCCACGACGACTTGCACCTGGAGGGTGGATGAAGGTGCCCAGG 4440
QY 4459 AATAACAAAGGAGCAGAAAGCTGGGACAGGAAGTACATTGCTCGAGGGATCAAAA 4518
Db 4441 AATAACAAAGGAGCAGAAAGCTGGGACAGGAAGTACATTGCTCGAGGGATCAAAA 4500
QY 4519 GTCCCTCATTTATGCAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAATTTGAG 4578
Db 4501 GTCCCTCATTTATGCAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAATTTGAG 4560
QY 4579 CTGTGCTTCCCGACGGGAGTGTATCTATTCAATGTTGCCGTTGGTCTTCCGAATCCGCA 4638
Db 4561 CTGTGCTTCCCGACGGGAGTGTATCTATTCAATGTTGCCGTTGGTCTTCCGAATCCGCA 4620
QY 4639 AATAACGCAAGCAGAAAGCAGAGAGTGTATTAATCTGTTGGAACCTCCCTGCTG 4698
Db 4621 AATAACGCAAGCAGAAAGCAGAGAGTGTATTAATCTGTTGGAACCTCCCTGCTG 4680

QY 4599 AAACCTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTGCCCTTCAGTGACCAGGTG 4758
Db 4581 AAACCTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTGCCCTTCAGTGACCAGGTG 4740
QY 4759 GTGTTGTGGGCAACCGAGGAAGGCTCTACGCCCTGAATGCTTTGAAAAACTCCCTTAACC 4818
Db 4741 GTGTTGTGGGCAACCGAGGAAGGCTCTACGCCCTGAATGCTTTGAAAAACTCCCTTAACC 4800
QY 4819 CATCTCCACAGAAATTGAGCAGCTTCCAAATTTATATTATCAAGGACCTGGAGAAAGCTA 4878
Db 4801 CATCTCCACAGAAATTGAGCAGCTTCCAAATTTATATTATCAAGGACCTGGAGAAAGCTA 4860
QY 4879 CTCATGATAGCAGGAGAAAGCGGCACTGTGCTTGTGGACGTTGAAGAAAGTGAACAG 4938
Db 4861 CTCATGATAGCAGGAGAAAGCGGCACTGTGCTTGTGGACGTTGAAGAAAGTGAACAG 4920
QY 4939 TCCCTGSCCAGTCCCACTTGCCTGCCAGCCGAGCATCTCACCCAACTTTTGAAGCT 4998
Db 4921 TCCCTGSCCAGTCCCACTTGCCTGCCAGCCGAGCATCTCACCCAACTTTTGAAGCT 4980
QY 4999 GTCAAGGCTGCCACTTGTGTTGGGGCAGGCAAGATTGAGAAACGGGCTCTGCATCTGTGCA 5058
Db 4981 GTCAAGGCTGCCACTTGTGTTGGGGCAGGCAAGATTGAGAAACGGGCTCTGCATCTGTGCA 5040
QY 5059 GCCATGCCAGCAAAAGTGTCTATTTCTCCGCTACAAACGAAAACTTCAGCAAAATACTGATC 5118
Db 5041 GCCATGCCAGCAAAAGTGTCTATTTCTCCGCTACAAACGAAAACTTCAGCAAAATACTGATC 5100
QY 5119 CGGAAAGAGATAGAGACTCAGAGCCCTGCAGCTGTATTCACCTTCACCAATTTACAGTATC 5178
Db 5101 CGGAAAGAGATAGAGACTCAGAGCCCTGCAGCTGTATTCACCTTCACCAATTTACAGTATC 5160
QY 5179 CTCATTGGAACCAATAAATTCTAGAAATCGACATGAAGCAGTACACGCTCGAGGAATTC 5238
Db 5161 CTCATTGGAACCAATAAATTCTAGAAATCGACATGAAGCAGTACACGCTCGAGGAATTC 5220
QY 5239 CTGATTAAGATGACCAATTCCTTGGCACCTGCTGTGTTTCCGCTCTTCCACAGCTTC 5298
Db 5221 CTGATTAAGATGACCAATTCCTTGGCACCTGCTGTGTTTCCGCTCTTCCACAGCTTC 5280
QY 5299 CCTGTCTCAATCTGTGAGGTGAACAGCGCAGGGCAGGACGAGGAGTACTTGTCTGTGTTTC 5358
Db 5281 CCTGTCTCAATCTGTGAGGTGAACAGCGCAGGGCAGGACGAGGAGTACTTGTCTGTGTTTC 5340
QY 5359 CACGAATTTGGAGTGTCTGTGGATTCTTACGGAAGACGTAGCCGACAGACGATCTCAAG 5418
Db 5341 CACGAATTTGGAGTGTCTGTGGATTCTTACGGAAGACGTAGCCGACAGACGATCTCAAG 5400
QY 5419 TGGAGTCCGCTTACCTTTGGCCTTTGCCCTACAGAGAACCTATCTGTTGTGACCCACTTC 5478
Db 5401 TGGAGTCCGCTTACCTTTGGCCTTTGCCCTACAGAGAACCTATCTGTTGTGACCCACTTC 5460
QY 5479 AACTCACTCGAAGTAAATTGAGATCCAGGACGCTTCTCAGACGGAACCCCTGCCGAGCG 5538
Db 5461 AACTCACTCGAAGTAAATTGAGATCCAGGACGCTTCTCAGACGGAACCCCTGCCGAGCG 5520
QY 5539 TACCTGGACATCCGGAACCCGCTACCTGGGCCCTGCCATTTCTCAGAGCGATTTTAC 5598
Db 5521 TACCTGGACATCCGGAACCCGCTACCTGGGCCCTGCCATTTCTCAGAGCGATTTTAC 5580
QY 5599 TTGCGCTCCTCATACAGGATAAATTAAAGGTCAATTTGCTGCAAGGAAACCTCGTGAAG 5658
Db 5581 TTGCGCTCCTCATACAGGATAAATTAAAGGTCAATTTGCTGCAAGGAAACCTCGTGAAG 5640
QY 5659 GAGTCCGCACTGNAACACCGGCGCTTCCACCTCCCGCAGCAGCCCAACAGGGA 5718
Db 5641 GAGTCCGCACTGNAACACCGGCGCTTCCACCTCCCGCAGCAGCCCAACAGGGA 5700
QY 5719 GGCCCAACCCAGTACAAACGAGCAGCATCAACAAAGCGCTGCGCTCCAGCCCGCGCCGCC 5778
Db 5701 GGCCCAACCCAGTACAAACGAGCAGCATCAACAAAGCGCTGCGCTCCAGCCCGCGCCGCC 5760
QY 5779 GAAGGCCCCAGCCACCCCGAGAGCCCAAGCACACCCCAACCGCTACCCGAGAGGGCGGACC 5838

5761	Db	GAAGGCCCCAGCCACCCGGAGAGCCAAAGCACACCCACCGCTACCCGAGGGCGCGACC	5820
5839	Qy	GAGCTGGCAGGACAAAGTCTCTCTGGCCGCCCTTGAGCGACAGAAAGTCTCCCGGCCCGG	5898
5821	Db	GAGCTGGCAGGACAAAGTCTCTGGCCGCCCTTGAGCGACAGAAAGTCTCCCGGCCCGG	5880
5899	Qy	ATGCTCAGCACCGGAGAGAGCGGTCCTCCCGGAGAGCTGTTTGAAGACAGACAGCGGGC	5958
5881	Db	ATGCTCAGCACCGGAGAGAGCGGTCCTCCCGGAGAGCTGTTTGAAGACAGACAGCGGGC	5940
5959	Qy	CGGCTGCTCGGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAAGGGAACAGAGGCGAG	6018
5941	Db	CGGCTGCTCGGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAAGGGAACAGAGGCGAG	6000
6019	Qy	AGTGCCCTCAAGTTTTCACGGTTAAACATGTGCACCTATTATGACTGGAAATAAAAGCTG	6078
6001	Db	AGTGCCCTCAAGTTTTCACGGTTAAACATGTGCACCTATTATGACTGGAAATAAAAGCTG	6060
6079	Qy	GACAACTGCCAGCTAACTGGTCAGTCCGTGAGGATCATCCAGCTGAATCGAGAAATCCCG	6138
6061	Db	GACAACTGCCAGCTAACTGGTCAGTCCGTGAGGATCATCCAGCTGAATCGAGAAATCCCG	6120
6139	Qy	CACAGGTTGAAAGTCTGTTCTGAGAACAGATTATTGC	6177
6121	Db	CACAGGTTGAAAGTCTGTTCTGAGAACAGATTATTGC	6159

RESULT 5
IIS-10-415-011-43

```

US/10-415-011-43
Sequence 43, Application US/10415011
Publication No. US20040053394A1
GENERAL INFORMATION:
APPLICANT: INCYTE CORPORATION
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: BAUGHN, Mariar R.
APPLICANT: CHAWLA, Narinder K.
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: XU, Yuming
APPLICANT: ARVIZU, Chandra S.
APPLICANT: YAO, Monique G.
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: DING, Li
APPLICANT: TANG, Y. Tom
APPLICANT: HAPALIA, April J.A.
APPLICANT: NGUYEN, Dannel B.
APPLICANT: GANDHI, Ameena R.
APPLICANT: YUE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: BANDMAN, Olga
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: LAL, Preeti G.
APPLICANT: RECIPON, Shirley A.
APPLICANT: LU, Dying Aina M.
APPLICANT: BOROWSKY, Mark L.
APPLICANT: THORNTON, Michael B.
APPLICANT: SWARNAKER Anita
APPLICANT: THIANGAVELU, Kavitha
APPLICANT: KHAN, Farrah A.
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: HUMAN KINASES
FILE REFERENCE: PT-0262 USN
CURRENT APPLICATION NUMBER: US/10/415,011,011
CURRENT FILING DATE: 2003-04-18
PRIOR APPLICATION NUMBER: PCT/US01/47728
PRIOR FILING DATE: 2001-10-20
PRIOR APPLICATION NUMBER: US 60/242,410
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/244,068
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60/245,708
PRIOR FILING DATE: 2000-11-03

```

```

; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 43
; LENGTH: 6298
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7484498CB1
US-10-415-011-43

Query Match      86.5%; Score 5684.4; DB 13; Length 6298;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5979; Conservative 0; Mismatches 36; Indels 243; Gaps 3;

QY 1 AGAGCGCCAGTGGGAGATGTTGAAGTTCAAATATGGAGCGGGAATCTTTGGATGCT 60
DB 37 AGAGCGCCAGTGGGAGATGTTGAAGTTCAAATATGGAGCGGGAATCTTTGGATGCT 96
QY 61 GFTGCTGTGAACCAATTGCCAGCCGGGCTCCAGGCTGAATCTGTCTTCAGGGGAAA 120
DB 97 GFTGCTGTGAACCAATTGCCAGCCGGGCTCCAGGCTGAATCTGTCTTCAGGGGAAA 156
QY 121 CCACCCCTTATGACTCAAACAGACAGATGTCTCTCTTCCCGAGAAAGGATATTAGATGCC 180
DB 157 CCACCCCTTATGACTCAAACAGACAGATGTCTCTCTTCCCGAGAAAGGATATTAGATGCC 216
QY 181 CTCTTTGTCTCTTTGAAAGATGCTAGCTAGCTGTCTCTGATGAAGATTAAAGCACGTGAGC 240
DB 217 CTCTTTGTCTCTTTGAAAGATGCTAGCTAGCTGTCTCTGATGAAGATTAAAGCACGTGAGC 276
QY 241 AACTTTGTCCGGAAGTATTTCCGCACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCA 300
DB 277 AACTTTGTCCGGAAGTATTTCCGCACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCA 336
QY 301 AAGGACTTCGAAGTCAGAAAGTCTTTGAGTTGTGGTCACTTTGCTGAAGTCAGGTGGTA 360
DB 337 AAGGACTTCGAAGTCAGAAAGTCTTTGAGTTGTGGTCACTTTGCTGAAGTCAGGTGGTA 396
QY 361 AGAGAGAAACAAACCGGGGACATCTATGCTATGAAGTATGAGAGAAAGGCTTTATTG 420
DB 397 AGAGAGAAACAAACCGGGGACATCTATGCTATGAAGTATGAGAGAAAGGCTTTATTG 456
QY 421 GCCCAGGACAGGTTTCATTTTTTGGAGGAGCGGAAACATATTATCTCGAACCAAGC 480
DB 457 GCCCAGGACAGGTTTCATTTTTTGGAGGAGCGGAAACATATTATCTCGAACCAAGC 516
QY 481 CCGTGGATCCCCAATTAAGTATGCTTTTCAGGACAAATAATCACTTTATCTGATGGAG 540
DB 517 CCGTGGATCCCCAATTAAGTATGCTTTTCAGGACAAATAATCACTTTATCTGATGGAG 576
QY 541 GAATATCAGCTGGAGGAGCTTGCTGCATCTTTTGAATAGATATGAGACCAAGTTAGAT 600
DB 577 GAATATCAGCTGGAGGAGCTTGCTGCATCTTTTGAATAGATATGAGACCAAGTTAGAT 636
QY 601 GAAACCTGATACAGTTTTTACCTAGCTAGCTGAATTTTGGCTGTTCACAGCGTTCACTG 660
DB 637 GAAACCTGATACAGTTTTTACCTAGCTAGCTGAATTTTGGCTGTTCACAGCGTTCACTG 696
QY 661 ATGGGATACGTGCATCGACATCAAGCTCGAACAATTCCTGTTGACCGCACAGACAC 720
DB 697 ATGGGATACGTGCATCGACATCAAGCTCGAACAATTCCTGTTGACCGCACAGACAC 756
QY 721 ATCAAGCTGGTGATTTTGGATCTGCGCGGAAAAATGAATTTCAAAACAAGATGGTGAATGCC 780

```


Qy	2893	CAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCAAGACCTGGAGGAGCAG	2952
Db	2977	CAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCAAGACCTGGAGGAGCAG	3036
Qy	2953	CTAAACAGCTGACCGGAGCAACGCTGAACTCAACAAACCAAACTTCTACTTTGTCCAAA	3012
Db	3037	CTAAACAGCTGACCGGAGCAACGCTGAACTCAACAAACCAAACTTCTACTTTGTCCAAA	3096
Qy	3013	CAACTCGATGAGGCTTCTGGCGCAACCGACGAGATTCTACAACCTGCGAAGTGAAGTGAC	3072
Db	3097	CAACTCGATGAGGCTTCTGGCGCAACCGACGAGATTCTACAACCTGCGAAGTGAAGTGAC	3156
Qy	3073	CATCTCGCGCGGAGATCAAGACGAGAGATCGAGCTTACCAGCCAGACGCAAAACGATG	3132
Db	3157	CATCTCGCGCGGAGATCAAGACGAGAGATCGAGCTTACCAGCCAGACGCAAAACGATG	3216
Qy	3133	GAGGCTCTGAAGACCAACGTGCACCATGCTGGAGGAAACAGTTCATGGATTTTGGAGGCCCTA	3192
Db	3217	GAGGCTCTGAAGACCAACGTGCACCATGCTGGAGGAAACAGTTCATGGATTTTGGAGGCCCTA	3276
Qy	3193	AACGATGAGCTGCTAGAAAAAGAGCGCAGTGGAGGCCCTGGAGAGCGTCTCTGGGTGAT	3252
Db	3277	AACGATGAGCTGCTAGAAAAAGAGCGCAGTGGAGGCCCTGGAGAGCGTCTCTGGGTGAT	3336
Qy	3253	GAGAAATCCCAAGTTTGAGTGTTCGGCTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAA	3312
Db	3337	GAGAAATCCCAAGTTTGAGTGTTCGGCTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAA	3396
Qy	3313	CAGAGCAGGCGAGAGCCGATCAGCGATCAACGAGTCTCCGAGGTGGAGCTGGCA	3372
Db	3397	CAGAGCAGGCGAGAGCCGATCAGCGATCAACGAGTCTCCGAGGTGGAGCTGGCA	3456
Qy	3373	GTGAAGAGACACAAGGCTGAGATTTCGCTCTGCAGCAGCGCTCTCAAAGAGCAGAAGCTG	3432
Db	3457	GTGAAGAGACACAAGGCTGAGATTTCGCTCTGCAGCAGCGCTCTCAAAGAGCAGAAGCTG	3516
Qy	3433	AAGCCCGAGAGCCCTCTCTGCAAGCTCAATGACCTGGAGAGAAGCAATGCTATGCTTGAA	3492
Db	3517	AAGCCCGAGAGCCCTCTCTGCAAGCTCAATGACCTGGAGAGAAGCAATGCTATGCTTGAA	3576
Qy	3493	ATGAAATCCCGAGCTTACAGCAGAGCTCGAGACTGAAACGAGCTCAACAGAGGCTT	3552
Db	3577	ATGAAATCCCGAGCTTACAGCAGAGCTCGAGACTGAAACGAGCTCAACAGAGGCTT	3636
Qy	3553	CTGGAAGAGCAAGCCAAATTTACAGCAGCAGATGGAACCTGCAGAAAAATCACATTTTCCGT	3612
Db	3637	CTGGAAGAGCAAGCCAAATTTACAGCAGCAGATGGAACCTGCAGAAAAATCACATTTTCCGT	3696
Qy	3613	CTGACTCAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAAAGT	3672
Db	3697	CTGACTCAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAAAGT	3756
Qy	3673	GACTTGAGTATCAGCTGGAAAAACATTCAGGTCTCTATTCTCATGAAAAGGTGAAAAATG	3732
Db	3757	GACTTGAGTATCAGCTGGAAAAACATTCAGGTCTCTATTCTCATGAAAAGGTGAAAAATG	3816
Qy	3733	GAAGGCACTATTCTCAACAAAACAAACTCATTTGATTTTCTGAAGCCAAAAATGACACAA	3792
Db	3817	GAAGGCACTATTCTCAACAAAACAAACTCATTTGATTTTCTGAAGCCAAAAATGACACAA	3876
Qy	3793	CTTGCTAAAAAGAAAAAGGGTTTATTTAGTCGACGGAAAGAGGACCCTGCTTTACCCACA	3852
Db	3877	CTTGCTAAAAAGAAAAAGGGTTTATTTAGTCGACGGAAAGAGGACCCTGCTTTACCCACA	3892
Qy	3853	CAGGTTCCCTCTGCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCTCGCTGT	3912
Db	3893	-AGGTTCCCTCTGCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCTCGCTGT	3951
Qy	3913	GCAGAGCTAGAGNAGCCCTTCAGAGACCCGATCGAGCTCCGCTCCGCTCCGCGGAGAA	3972
Db	3952	GCAGAGCTAGAGNAGCCCTTCAGAGACCCGATCGAGCTCCGCTCCGCTCCGCGGAGAA	4011
Qy	3973	GCTGCCACCGCAAAAGCAACGAGACCCACACCCATCCACGCCAGCCACCGCGAGCAG	4032

[illegible]

QY 2791 GCCTTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGAGCTGGAGAGACACACAGAGAGCT 2850
Db 2821 GCCTTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGAGCTGGAGAGACACACAGAGAGCT 2880
QY 2851 GAAGAGAGATCCAGAGCTCAGCGACATAGAGATGAATCCAGCGCAATTTGATGCT 2910
Db 2881 GAAGAGAGATCCAGAGCTCAGCGACATAGAGATGAATCCAGCGCAATTTGATGCT 2940
QY 2911 CTTCTGAACAGCTGTAATCTAATCAAGAGCTGGAGAGAGCTAAACAGCTGACCGAG 2970
Db 2941 CTTCTGAACAGCTGTAATCTAATCAAGAGCTGGAGAGAGCTAAACAGCTGACCGAG 3000
QY 2971 GACAAAGCTGAATCAACAAACAAACTTCTACTTGTCCAAACAACCTCGATGAGCTTCT 3030
Db 3001 GACAAAGCTGAATCAACAAACAAACTTCTACTTGTCCAAACAACCTCGATGAGCTTCT 3060
QY 3031 GCGCCCAACGAGAGATTTGATCAAACTGCGAAGTGAAGTGACCATCTCCGCGGAGATC 3090
Db 3061 GCGCCCAACGAGAGATTTGATCAAACTGCGAAGTGAAGTGACCATCTCCGCGGAGATC 3120
QY 3091 ACGGAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGAGGCTCTGAAGACCAAG 3150
Db 3121 ACGGAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGAGGCTCTGAAGACCAAG 3180
QY 3151 TGCACCATCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210
Db 3181 TGCACCATCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240
QY 3211 AAAGAGCGCAGTGGAGGCCCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3270
Db 3241 AAAGAGCGCAGTGGAGGCCCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
QY 3271 TGTCCGGTTTCGAGAGCTGCAGAGATGTCTGACACCGAGAAACAGAGAGGCGAGAGCC 3330
Db 3301 TGTCCGGTTTCGAGAGCTGCAGAGATGTCTGACACCGAGAAACAGAGAGGCGAGAGCC 3360
QY 3331 GATCAGCGATCACCGAGTCTCCAGGTGTGGAGCTGGCGAGTGAAGAGCACAGGCT 3390
Db 3361 GATCAGCGATCACCGAGTCTCCAGGTGTGGAGCTGGCGAGTGAAGAGCACAGGCT 3420
QY 3391 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGAGCTCAAGGCCGAGAGCCCTCTCT 3450
Db 3421 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGAGCTCAAGGCCGAGAGCCCTCTCT 3480
QY 3451 GACAGCTCAATGACTGAGAGAGAGCATGTATGCTTGAATGAATGCCGAGAGCTTA 3510
Db 3481 GACAGCTCAATGACTGAGAGAGAGCATGTATGCTTGAATGAATGCCGAGAGCTTA 3540
QY 3511 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3570
Db 3541 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3600
QY 3571 TTACAGCAGAGAGTGAACCTGAGAGAAATCAATTTTCCGCTCTGACTCAAGGACTGCAA 3630
Db 3601 TTACAGCAGAGAGTGAACCTGAGAGAAATCAATTTTCCGCTCTGACTCAAGGACTGCAA 3660
QY 3631 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAGAGAGTGAATGAGATCAGCTG 3690
Db 3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAGAGAGTGAATGAGATCAGCTG 3720
QY 3691 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAATGGAAGGCACTATTCTCAA 3750
Db 3721 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAATGGAAGGCACTATTCTCAA 3780
QY 3751 CAAAACAAATCATTTGATTTCTGACGCAAAATGGAACCACTGCTTAAAGAAAAAG 3810
Db 3781 CAAAACAAATCATTTGATTTCTGACGCAAAATGGAACCACTGCTTAAAGAAAAAG - - 3838
QY 3811 GCTTTATTTAGTCGAGGGAAGAGGACCTGCTTTTACCACACAGGTTCTCTCAGTAC 3870
Db 3839 -----AGGTTCTCTCAGTAC 3855

QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGAGCTCGCTGTGCAGAGCTAGAGAGCC 3930
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGAGCTCGCTGTGCAGAGCTAGAGAGCC 3915
QY 3931 CTTTCAAGAGACCCGATCGAGCTCCGGTCCGCGCGGAGGAAGTGCCTCCACCGCAAGCA 3990
Db 3916 CTTTCAAGAGACCCGATCGAGCTCCGGTCCGCGCGGAGGAAGTGCCTCCACCGCAAGCA 3975
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCAGCTGAGCCCTGCGGCCCGCCATTCGCTG 4050
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCCAGCTGAGCCCTGCGGCCCGCCATTCGCTG 4035
QY 4051 ATCTGCGGCTCGCCAGAGCACCAGCCAGTGCATGAGCCCTGCGGCCCGCCATTCGCTG 4110
Db 4036 ATCTGCGGCTCGCCAGAGCACCAGCCAGTGCATGAGCCCTGCGGCCCGCCATTCGCTG 4095
QY 4111 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCCGCGCTTTAAGGAACCGATGCAC 4170
Db 4096 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCCGCGCTTTAAGGAACCGATGCAC 4155
QY 4171 CACAATATTTCTCACCAGTTCAAGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4230
Db 4156 CACAATATTTCTCACCAGTTCAAGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4215
QY 4231 TGTCTGATACCGTGCATTTTGGACGCCAGGCATCCAAATGTCTCGAATCTCAGGTGATG 4290
Db 4216 TGTCTGATACCGTGCATTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCTCAGGTGATG 4275
QY 4291 TGTCAACCCCAAGTCTCCAGTGTCTTCCAGCCAGCTGCGGCTTGCCTGTGATATGCC 4350
Db 4276 TGTCAACCCCAAGTCTCCAGTGTCTTCCAGCCAGCTGCGGCTTGCCTGTGATATGCC 4335
QY 4351 ACACATTTCAACCCAGGCTTCTGCGGTGACAAATGAATCTCCAGGCTTCCAGACCAAG 4410
Db 4336 ACACATTTCAACCCAGGCTTCTGCGGTGACAAATGAATCTCCAGGCTTCCAGACCAAG 4395
QY 4411 GAGCCAGCAGAGCTTGCACCTTGGAAAGGTGATGAAGTGCCTCCAGGAATAACAAACGA 4470
Db 4396 GAGCCAGCAGAGCTTGCACCTTGGAAAGGTGATGAAGTGCCTCCAGGAATAACAAACGA 4455
QY 4471 GGACAGCAAGCTGGGACAGAGAGTACATTTGCTGAGGGATCAAAAGTCTCATTTAT 4530
Db 4456 GGACAGCAAGCTGGGACAGAGAGTACATTTGCTGAGGGATCAAAAGTCTCATTTAT 4515
QY 4531 GACAAATGAACCCAGAGAGCTGACAGAGCCCGGTGGAAGAATTTGAGTGTGCTTCCC 4590
Db 4516 GACAAATGAACCCAGAGAGCTGACAGAGCCCGGTGGAAGAATTTGAGTGTGCTTCCC 4575
QY 4591 GACGGGATCTATCTATTCATGCTGCGTGTGCTTCCGAACTCGCAATACAGCCAAA 4650
Db 4576 GACGGGATCTATCTATTCATGCTGCGTGTGCTTCCGAACTCGCAATACAGCCAAA 4635
QY 4651 GCA----- 4653
Db 4636 GCAGATGTCCCATACATACACTGAAGATGAATCTCACCCGACACACCTGCTGCGCCGGG 4695
QY 4654 ----- 4653
Db 4696 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACGCTTA 4755
QY 4654 -----GAAAAAGCAGAACTGATGCTAAACTG 4680
Db 4756 GAATCAGTTGTCCAGGTGGAGAGTTCTAGGAAAAAAGCAGAGCTGATGCTAACTG 4815
QY 4681 CTTGGAAAATCTCCCTGCTGAAAACCTGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4740
Db 4816 CTTGGAAAATCTCCCTGCTGAAAACCTGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4875
QY 4741 CCCTTCAGTGACAGAGTGTGTTGTTGGCCACCGAGGAAGGGCTCTACGCCCTGATATGC 4800
Db 4876 CCCTTCAGTGACAGAGTGTGTTGTTGGCCACCGAGGAAGGGCTCTACGCCCTGATATGC 4935
QY 4801 TTGAAAAATCTCCCTAAACCCATGTCCAGGAATTTGGAGCAGCTTTTCCAAATTTATATTATC 4860

Db 4936 TTGAAAATCTCCCTAACCCATGTGTCAGGAATTTGGAGCAGTCTTCCAAATTTATATATC 4995
QY 4861 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTGTCTTGTGGAC 4920
Db 4996 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTGTCTTGTGGAC 5055
QY 4921 GTCAAGAAAGTGAACAGTCCCTGGCCAGTCCCAGTCCCTGCGCCAGCCGACATCTCA 4980
Db 5056 GTCAAGAAAGTGAACAGTCCCTGGCCAGTCCCAGTCCCTGCGCCAGCCGACATCTCA 5115
QY 4981 CCCAAATTTTGAAGCTGTCAAGGCTGCCATTTGTTGGGCGAGGCAAGATTGAGAAC 5040
Db 5116 CCCAAATTTTGAAGCTGTCAAGGCTGCCATTTGTTGGGCGAGGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTGTCTTCTCGCTTACACGAAAC 5100
Db 5176 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTGTCTTCTCGCTTACACGAAAC 5235
QY 5101 CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTCAGCTGTATCCAC 5160
Db 5236 CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTCAGCTGTATCCAC 5295
QY 5161 TTCACCAATACAGTATCTCTTGAACCAATAATTTCTAGAAATCGACATGAAGCAG 5220
Db 5296 TTCACCAATACAGTATCTCTTGAACCAATAATTTCTAGAAATCGACATGAAGCAG 5355
QY 5221 TAGACGCTCAGGAATTCCTCGATAGATGACCATTCCTTGGCAGCTGTGTGTGCC 5280
Db 5356 TAGACGCTCAGGAATTCCTCGATAGATGACCATTCCTTGGCAGCTGTGTGTGCC 5415
QY 5281 GCCTCTTCAACAGCTTCCTCTGTCTCAATCGTGCAGTGAACAGCGCAGGCGAGGAG 5340
Db 5416 GCCTCTTCAACAGCTTCCTCTGTCTCAATCGTGCAGTGAACAGCGCAGGCGAGGAG 5475
QY 5341 GAGTACTGTGTGTGTTCACCAATTTGGAGTGTTCGTGGATTTTACGGAAGAGCTAGC 5400
Db 5476 GAGTACTGTGTGTGTTCACCAATTTGGAGTGTTCGTGGATTTTACGGAAGAGCTAGC 5535
QY 5401 CGCACAGACATCTCAAGTGGAGTCCCTTACCTTGGCCCTTGGCTTACAGAGAACCTAT 5460
Db 5536 CGCACAGACATCTCAAGTGGAGTCCCTTACCTTGGCCCTTGGCTTACAGAGAACCTAT 5595
QY 5461 CTGTTGTGACCCACTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5520
Db 5596 CTGTTGTGACCCACTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5655
QY 5521 GGGACCCCTCCGAGGTTACCTGGACATCCCGAACCCCGGCTACCTGGCCCTGCCATT 5580
Db 5656 GGGACCCCTCCGAGGTTACCTGGACATCCCGAACCCCGGCTACCTGGCCCTGCCATT 5715
QY 5581 TCCTCAGAGCGGATTTACTTGGGCTCCTATACAGGATAAATTAAGGTCATTTGCTGC 5640
Db 5716 TCCTCAGAGCGGATTTACTTGGGCTCCTATACAGGATAAATTAAGGTCATTTGCTGC 5775
QY 5641 AAGGGAACTCTGTGAAGAGTCCGGACATGAACCAACCGGGGCCCTGCACCTCCCGC 5700
Db 5776 AAGGGAACTCTGTGAAGAGTCCGGACATGAACCAACCGGGGCCCTGCACCTCCCGC 5835
QY 5701 AGCAGCCCCAACAGCGAGGCCACCCAGCTACACGAGCACATCAACGAGCGGTGGCC 5760
Db 5836 AGCAGCCCCAACAGCGAGGCCACCCAGCTACACGAGCACATCAACGAGCGGTGGCC 5895
QY 5761 TCCAGCCCCAGCGCCGAGGCCACCCAGCTACCCCGAGAGCAAGCAACCCACCCG 5820
Db 5896 TCCAGCCCCAGCGCCGAGGCCACCCAGCTACCCCGAGAGCAAGCAACCCACCCG 5955
QY 5821 TACCGCGAGGGCGGACCGAGCTGCCAGGACAAGTCTCTGGCCGCCCTCGAGGCA 5880
Db 5956 TACCGCGAGGGCGGACCGAGCTGCCAGGACAAGTCTCTGGCCGCCCTCGAGGCA 6015
QY 5881 GAGAAGTCTCCCGCGGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGGCTGT 5940

Db 6016 GAGAGTCCCCCGCGGATACTCAGACCGCGGAGAGCGGTCCCCCGGAGGCTGTTT 6075
QY 5941 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCTCCAGGTG 6000
Db 6076 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCTCCAGGTG 6135
QY 6001 AACAGGGAAGAGGCGCAGAGTGC 6023
Db 6136 AACAGGCTTGGGACCAAGTCTTC 6158
RESULT 7
US-10-262-511-1
; Sequence 1, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Beighs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curasequist version 0.1
; SEQ ID NO 1

```
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6159)
US-10-262-511-1

Query Match      85.7%; Score 5631; DB 13; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;

QY 19 ATGTTGAAGTCAAAATATGAGCGGGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT 78
Db 1 ATGTTGAAGTCAAAATATGAGCGGGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT 60

QY 79 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCTTTATGACTCAA 138
Db 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCTTTATGACTCAA 120

QY 139 CAGCAGATGTCCTCTCTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198
Db 121 CAGCAGATGTCCTCTCTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180

QY 199 GAATGAGTCAACCTGCTCTGTATGAAGATTAAAGCACTGAGCAACTTTGTCGGGAAGTAT 258
Db 181 GAATGAGTCAACCTGCTCTGTATGAAGATTAAAGCACTGAGCAACTTTGTCGGGAAGTAT 240

QY 259 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCCTTCGGCAAGGACTTCGAACTGAGA 318
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCCTTCGGCAAGGACTTCGAACTGAGA 300

QY 319 AGTCTTTGAGGTGTGGTCACTTTCTGTAAGTGCAGGTGGTAAAGAGAGAAACACCGGG 378
Db 301 AGTCTTTGAGGTGTGGTCACTTTCTGTAAGTGCAGGTGGTAAAGAGAGAAACACCGGG 360

QY 379 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 438
Db 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 420

QY 439 TTTTGTGAGGAGAGCGGAAACATATATCTCGAAGCACAAGCCGTGGATCCCCCAATTA 498
Db 421 TTTTGTGAGGAGAGCGGAAACATATATCTCGAAGCACAAGCCGTGGATCCCCCAATTA 480

QY 499 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 558
Db 481 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540

QY 559 GACTTGTCTGCTACTTTTGAATAGATATGAGGACAGTTAGATGAAACCTGATACAGTTT 618
Db 541 GACTTGTCTGCTACTTTTGAATAGATATGAGGACAGTTAGATGAAACCTGATACAGTTT 600

QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTCAACAGCTTCATCTGATGGGATACGTGCATCGA 678
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAACAGCTTCATCTGATGGGATACGTGCATCGG 660

QY 679 GACATCAAGCCCTGAGAACATCTCTTGAACGACAGGACACATCAAGCTGGTGAATTTT 738
Db 661 GACATCAAGCCCTGAGAACATCTCTTGAACGACAGGACACATCAAGCTGGTGAATTTT 720

QY 739 GGATCTGCCCGGAAAATGAATTCAAACAAAGATGGTGAATGCCAACTCCCGATTGGGACC 798
Db 721 GGATCTGCCCGGAAAATGAATTCAAACAAAGATGGTGAATGCCAACTCCCGATTGGGACC 777

QY 799 CCAGATTATCATGGCTCTGTAAGTGTGATGTAAGACGGGGATGGAAGGACACCTTAC 858
Db 778 CCAGATTATCATGGCTCTGTAAGTGTGATGTAAGACGGGGATGGAAGGACACCTTAC 837

QY 859 GSCCTGGAGTGTGACTGGTGTGCTGGGCGGTGATGCTGATGAGATGATTTATGGGAGA 918
Db 838 GSCCTGGAGTGTGACTGGTGTGCTGGGCGGTGATGCTGATGAGATGATTTATGGGAGA 897

QY 919 TCCCCCTTCGAGAGGGGAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 978
```

```
Db 898 TCCCCCTTCGAGAGGGGAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 957
QY 979 TTTTGTGAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1038
Db 958 TTTTGTGAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
QY 1039 TTGTTGTGCGGCGCAGAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCTTTCTTC 1098
Db 1018 TTGTTGTGCGGCGCAGAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCTTTCTTC 1077
QY 1099 TCTAAAAATTGACTGGAACAAACATTTCGTTAACTCTCTCCCTCCCTTCGTTCCACCCCTCAAG 1158
Db 1078 TCTAAAAATTGACTGGAACAAACATTTCGTTAACTCTCTCCCTCCCTTCGTTCCACCCCTCAAG 1137
QY 1159 TCTGACGATGACACCTCCAAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1218
Db 1138 TCTGACGATGACACCTCCAAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1197
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAGAACTGCGCTTTGTTGGGGTTTTCG 1278
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAGAACTGCGCTTTGTTGGGGTTTTCG 1257
QY 1279 TACAGCAAGCACTGGGATTTCTTTGTAGATCTGAGTCTGTTGTGCTGGGTTCTGGACTCC 1338
Db 1258 TACAGCAAGCACTGGGATTTCTTTGTAGATCTGAGTCTGTTGTGCTGGGTTCTGGACTCC 1317
QY 1339 CCTGCCAAGACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTACAGAC 1398
Db 1318 CCTGCCAAGACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTACAGAC 1377
QY 1399 TCTCAGACAAAGTGTTCACAAAGATGGAGCAGGAAATGACCCGGTTTACATCGGAGAGTGTCA 1458
Db 1378 TCTCAGACAAAGTGTTCACAAAGATGGAGCAGGAAATGACCCGGTTTACATCGGAGAGTGTCA 1437
QY 1459 GAGGTGAGGCTGTCTTAGTCAAGAGGAGTGGAGCTGAGGCTTGAAGCCCTTGAGACTCAGAGA 1518
Db 1438 GAGGTGAGGCTGTCTTAGTCAAGAGGAGTGGAGCTGAGGCTTGAAGCCCTTGAGACTCAGAGA 1497
QY 1519 TCCCTCTCGAGCAGGACCTTGTCTACTCATCACAGAAATGACAGTGTAGCTTAAAGCCGAAGT 1578
Db 1498 TCCCTCTCGAGCAGGACCTTGTCTACTCATCACAGAAATGACAGTGTAGCTTAAAGCCGAAGT 1557
QY 1579 TTGGAGCAAGCAACCGATGGAGGTGTCAGGAGGATGACAAAGCACTGACAGCTTCTCCAT 1638
Db 1558 TTGGAGCAAGCAACCGATGGAGGTGTCAGGAGGATGACAAAGCACTGACAGCTTCTCCAT 1617
QY 1639 GATATCAGAGCAGAGCCGAGGCTCCAGAGAAATCAAGAGCAGGAGTACAGGCTCAA 1698
Db 1618 GATATCAGAGCAGAGCCGAGGCTCCAGAGAACTCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677
QY 1699 GTGGAAGAAATGAGGTGATGATGAATCAAGTTTGAAGAGGATCTTGTCTCAGCAAGAA 1758
Db 1678 GTGGAAGAAATGAGGTGATGATGAATCAAGTTTGAAGAGGATCTTGTCTCAGCAAGAA 1737
QY 1759 CGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAGAAATTCAG 1818
Db 1738 CGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAGAAATTCAG 1797
QY 1819 CGGAAAGCAGACAAATGTTCAGCATAACTGTTGAAGCTTAAAGATCAAGGAGAGCTGAA 1878
Db 1798 CGGAAAGCAGACAAATGTTCAGCATAACTGTTGAAGCTTAAAGATCAGGGAAGCTGAA 1857
QY 1879 GTGGGAAATATCGGAAATCTGGAGAGAGTCAATGTCTGAGCAGAGCTCAAAATTCAGGAG 1938
Db 1858 GTGGGAAATATCGGAAATCTGGAGAGAGTCAATGTCTGAGCAGAGCTCAAAATTCAGGAG 1917
QY 1939 CTCCAAGAGAACTGGAGAGGCTG----- 1963
Db 1918 CTCCAAGAGAACTGGAGAGGCTGTAAGAGCCAGCAGAGGCTGAGAGCTGAGAGCTGAGAG 1977
QY 1964 -----CAAGAGGCGAGCGGAGGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAGAG 2010
```

Db 1978 AATATCCGCCAAGAGCGAGCGGAGCGGAGCTGGAGAGCTGCAGAACCCAGAG 2037
QY 2011 GATTCCTTCTGAAGGCATCAAGAAAGAGCTGGTGGAGAGCTGGAGAACCGCCGCAATCTCTG 2070
Db 2038 GATTCCTTCTGAAGGCATCAAGAAAGAGCTGGTGGAGAGCTGGAGAACCGCCGCAATCTCTG 2097
QY 2071 GAGAAACAAGGTAAAGAGACTAGAGACCATGGAGCGTGGAGAGAAACAGAGCTGAAGATGAC 2130
Db 2098 GAGAAACAAGGTAAAGAGACTAGAGACCATGGAGCGTGGAGAGAAACAGAGCTGAAGATGAC 2157
QY 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2190
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2217
QY 2191 AAACATCGGGAGGCCCAAGTCTCAGCCCGAGCAGCTAGAGTGCACCTGNAACAGAAAGAG 2250
Db 2218 AAACATCGGGAGGCCCAAGTCTCAGCCCGAGCAGCTAGAGTGCACCTGNAACAGAAAGAG 2277
QY 2251 CAGCACTATCAGGAAAGAGTTAAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2310
Db 2278 CAGCACTATCAGGAAAGAGTTAAAGTATTGGACATCAGATTAAGAGACCTGGCTGAC 2337
QY 2311 AAGGAGACACTGAGAAACATGATGACAGACAGGAGGAGGCGCCATGAGAAAGGCAAA 2370
Db 2338 AAGGAGACACTGAGAAACATGATGACAGACAGGAGGAGGCGCCATGAGAAAGGCAAA 2397
QY 2371 ATTCTCAGCGAAACAGAAAGCGATGATCAATGCTATGATTCCAGATTCAGATCCCTGGAA 2430
Db 2398 ATTCTCAGCGAAACAGAAAGCGATGATCAATGCTATGATTCCAGATTCAGATCCCTGGAA 2457
QY 2431 CAGAGGATTTGGAACTGCTGAAGCCATAAATTTGACGCAAAATAGCAGTCTTTTACC 2490
Db 2458 CAGAGGATTTGGAACTGCTGAAGCCATAAATTTGACGCAAAATAGCAGTCTTTTACC 2517
QY 2491 CAAGAGACATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2550
Db 2518 CAAGAGACATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2577
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2610
Db 2578 CTGGAGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2637
QY 2611 AAGATCAGCCACCAAGACACAGTGAACAAGATCGCTCTGGAATCTGGAGGAGCAGATTTG 2670
Db 2638 AAGATCAGCCACCAAGACACAGTGAACAAGATCGCTCTGGAATCTGGAGGAGCAGATTTG 2697
QY 2671 CGGGAGGTCACTAGACACAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2730
Db 2698 CGGGAGGTCACTAGACACAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2757
QY 2731 CTAAGCTCTCCCTGAGGAGCGAGTCAAGTTCAGAGCTTCAGAGCTGCAAGCGG 2790
Db 2758 CTAAGCTCTCCCTGAGGAGCGAGTCAAGTTCAGAGCTTCAGAGCTGCAAGCGG 2817
QY 2791 GCCTGGAGGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCAACAGAGCT 2850
Db 2818 GCCTGGAGGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCAACAGAGCT 2877
QY 2851 GAAGAGGAGATCAGGCACTTCAGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGT 2910
Db 2878 GAAGAGGAGATCAGGCACTTCAGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGT 2937
QY 2911 CTTCTGTAACAGCTGTAAGTATCAAGACCTGGAGGAGCTAAACAGCTGACCGAG 2970
Db 2938 CTTCTGTAACAGCTGTAAGTATCAAGACCTGGAGGAGCTAAACAGCTGACCGAG 2997
QY 2971 GACAAAGCTGAACTCAACAAACAAAATCTTCTATTGTCAAAACAACTCGATGAGGCTTCT 3030
Db 2998 GACAAAGCTGAACTCAACAAACAAAATCTTCTATTGTCAAAACAACTCGATGAGGCTTCT 3057
QY 3031 GCGGCCAAACGACAGATTTGACAACTGCGAAGTGAAGTGCACCTCCCGCGGAGATC 3090
Db 3058 GCGGCCAAACGACAGATTTGACAACTGCGAAGTGAAGTGCACCTCCCGCGGAGATC 3117

QY 3091 ACGGAACGAGATGCGAGCTTACAGCCAGAACCAACGATGAGGCTCTGAAGACCAAG 3150
Db 3118 ACGGAACGAGATGCGAGCTTACAGCCAGAACCAACGATGAGGCTCTGAAGACCAAG 3177
QY 3151 TGCAACCATGCTGGAGGAAACAGGTCAATGGATTTTGAGGCGCTTAAACGATGAGCTGTAGAA 3210
Db 3178 TGCAACCATGCTGGAGGAAACAGGTCAATGGATTTTGAGGCGCTTAAACGATGAGCTGTAGAA 3237
QY 3211 AAAGAGCGCAGTGGAGGCGCTGGAGGAGCTGCTGGGTGATGAGAAATCCCAAGTTTGA 3270
Db 3238 AAAGAGCGCAGTGGAGGCGCTGGAGGAGCTGCTGGGTGATGAGAAATCCCAAGTTTGA 3297
QY 3271 TGTCCGCTTCAGAGCTGCGAGAAATGCTGGAACCGAGAAACAGAGCAGGCGGAGAGCC 3330
Db 3298 TGTCCGCTTCAGAGCTGCGAGAAATGCTGGAACCGAGAAACAGAGCAGGCGGAGAGCC 3357
QY 3331 GATCAGCGGATCAACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGGAGCAAAAGCT 3390
Db 3358 GATCAGCGGATCAACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGGAGCAAAAGCT 3417
QY 3391 GAGATTCCTGCTTCGAGCAGGCTCTCAAGAGCAGAAAGCTGAAGCGCGAGAGCTCTCT 3450
Db 3418 GAGATTCCTGCTTCGAGCAGGCTCTCAAGAGCAGAAAGCTGAAGCGCGAGAGCTCTCT 3477
QY 3451 GACAAGCTCAATGACCTGGAGAAAGAGCTATGCTTGAATGAATGCTCCGAGGCTTA 3510
Db 3478 GACAAGCTCAATGACCTGGAGAAAGAGCTATGCTTGAATGAATGCTCCGAGGCTTA 3537
QY 3511 CAGCAGAGCTGGAGCTGAACGAGGCTCAACAGAGGCTCTGGAAGAGCAAGCAAA 3570
Db 3538 CAGCAGAGCTGGAGCTGAACGAGGCTCAACAGAGGCTCTGGAAGAGCAAGCAAA 3597
QY 3571 TTACAGCAGCAGATGGACCTGCGAGAAATCAATTTCCGCTCTGACTCAAGGAGCTCAA 3630
Db 3598 TTACAGCAGCAGATGGACCTGCGAGAAATCAATTTCCGCTCTGACTCAAGGAGCTCAA 3657
QY 3631 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTCACTGGAGTATCAGCTG 3690
Db 3658 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTCACTGGAGTATCAGCTG 3717
QY 3691 GAAACCAATTCAGGTTCTCTATCTCATGAAAGGTGAAATGGAAGCACTATTTCCTAA 3750
Db 3718 GAAACCAATTCAGGTTCTCTATCTCATGAAAGGTGAAATGGAAGCACTATTTCCTAA 3777
QY 3751 CAAACCAAACTCAATGATTTTCTGAAGCCAAATGGAACCACTGCTAATAAGAAAG 3810
Db 3778 CAAACCAAACTCAATGATTTTCTGAAGCCAAATGGAACCACTGCTAATAAGAAAG -- 3835
QY 3811 GGTTTATTTAGTCGACGGAAGAGGACCTGCTTTACCCACACAGGTTCTCTGCGAGTAC 3870
Db 3836 -----AGGTGCTCTGCGAGTAC 3852
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGTCAGAGCTTAGAGGAGCC 3930
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGTCAGAGCTTAGAGGAGCC 3912
QY 3931 CTTTCAGAGACCCGATCGAGCTCCGCTCCCGGAGGAGAACTGCCACCCGCAAGCA 3990
Db 3913 CTTTCAGAGACCCGATCGAGCTCCGCTCCCGGAGGAGAACTGCCACCCGCAAGCA 3972
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCAGCCAGGAGCAGAGATGCGCATGTGCGCC 4050
Db 3973 ACGGACCAACCAACCCATCCAGCCAGCCAGCCAGGAGCAGAGATGCGCATGTGCGCC 4032
QY 4051 ATCGTGGGTGCGCAGAGCAGCCAGTCCCATGAGCTGCTGGCCCGCCATCCAGC 4110
Db 4033 ATCGTGGGTGCGCAGAGCAGCCAGTCCCATGAGCTGCTGGCCCGCCATCCAGC 4092
QY 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTGGCGCTTTAAGGAACGATGAC 4170
Db 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTGGCGGTCTTAAGGAACGATGAC 4152


```

1 APPLICANT: Stone, David
2 APPLICANT: Gunther, Erik
3 APPLICANT: Ellerman, Karen
4 APPLICANT: Grosse, William M
5 APPLICANT: Alsbrook II, John P
6 APPLICANT: Lepley, Denise M
7 APPLICANT: Burgess, Catherine E
8 APPLICANT: Padigaru, Muralidhara
9 APPLICANT: Kekuda, Ramesh
10 APPLICANT: Spytek, Kimberly A
11 APPLICANT: Leach, Martin D
12 APPLICANT: Shimkets, Richard A
13 TITLE OF INVENTION: No. US20040043926A1el Proteins and Nucleic Acids Encoding Same
14 FILE REFERENCE: 21402-124
15 CURRENT APPLICATION NUMBER: US/09/964,956
16 CURRENT FILING DATE: 2001-09-26
17 PRIOR APPLICATION NUMBER: 60/235,631
18 PRIOR FILING DATE: 2000-09-27
19 PRIOR APPLICATION NUMBER: 60/235,633
20 PRIOR FILING DATE: 2000-09-27
21 PRIOR APPLICATION NUMBER: 60/235,808
22 PRIOR FILING DATE: 2000-09-27
23 PRIOR APPLICATION NUMBER: 60/236,064
24 PRIOR FILING DATE: 2000-09-27
25 PRIOR APPLICATION NUMBER: 60/236,065
26 PRIOR FILING DATE: 2000-09-27
27 PRIOR APPLICATION NUMBER: 60/236,066
28 PRIOR FILING DATE: 2000-09-27
29 PRIOR APPLICATION NUMBER: 60/236,135
30 PRIOR FILING DATE: 2000-09-28
31 PRIOR APPLICATION NUMBER: 60/237,434
32 PRIOR FILING DATE: 2000-10-03
33 PRIOR APPLICATION NUMBER: 60/238,321
34 PRIOR FILING DATE: 2000-10-05
35 PRIOR APPLICATION NUMBER: 60/238,399
36 PRIOR FILING DATE: 2000-10-06
37 PRIOR APPLICATION NUMBER: 60/238,396
38 PRIOR FILING DATE: 2000-10-06
39 PRIOR APPLICATION NUMBER: 60/276,667
40 PRIOR FILING DATE: 2001-03-16
41 PRIOR APPLICATION NUMBER: 60/294,823
42 PRIOR FILING DATE: 2001-05-31
43 PRIOR APPLICATION NUMBER: 60/304,868
44 PRIOR FILING DATE: 2001-07-12
45 NUMBER OF SEQ ID NOS: 127
46 SOFTWARE: PatentIn Ver. 2.1
47 SEQ ID NO 10
48 LENGTH: 6189
49 TYPE: DNA
50 ORGANISM: Homo sapiens
51 US-09-964-956-10

```

Query Match	85.7%;	Score 5631;	DB 13;	Length 6189;
Best Local Similarity	95.3%;	Pred. No. 0;		
Matches 5944;	Conservative 0;	Mismatches 45;	Indels 246;	Gaps 4;
Qy	19	ATGTTGAAGTTCAAATATGGAGCGCGAAATCCTTTGGATGCTGCTGCTGAACCCCAAT	78	
Db	1	ATGTTGAAGTTCAAATATGGAGCGCGNAATCCTTTGGATGCTGCTGCTGAACCCCAAT	60	
Qy	79	GCGAGCGGGCTCAGGCTGAATCTGTTCTTCAGGGGAAACACACCTTTATGACTCAA	138	
Db	61	GCGAGCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACACCTTTATGACTCAA	120	
Qy	139	CAGCAGATGTCCTCTCTTCCGAGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTTGAA	198	
Db	121	CAGCAGATGTCCTCTCTTCCGAGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTTGAA	180	
Qy	199	GAATGCAGTCAGCTGCTCTGATGAAGATTAAGCAGCTGACCACTTTGTCGGAAGTAT	258	
Db	181	GAATGCAGTCAGCTGCTCTGATGAAGATTAAGCAGCTGACCACTTTGTCGGAAGTGT	240	
Qy	259	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAGTCAGA	318	

Dd	1318	CCTGCCAAGACTAGCTCCATCGAAAGAAATCTTCATCAAAAGCAAGAGCTACAAGAC	1377
Qy	1399	TCTCAGGACAAGTGTCAACAAGATGGAGCAGGAGAAATGACCCGGTTACATCGGAGAGTGTCA	1458
Dd	1378	TCTCAGGACAAGTGTCAACAAGATGGAGCAGGAGAAATGACCCGGTTACATCGGAGAGTGTCA	1437
Qy	1459	GAGGTGGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1518
Dd	1438	GAGGTGGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1497
Qy	1519	TGCTCTCTGGAGAGGACCTTGCTACCTACATCAACAAGATGAGTAGCTTAAAGCGAAGT	1578
Dd	1498	TGCTCTCTGGAGAGGACCTTGCTACCTACATCAACAAGATGAGTAGCTTAAAGCGAAGT	1557
Qy	1579	TTGGACCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT	1638
Dd	1558	TTGGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT	1617
Qy	1639	GATATCAGAGACAGACCGGAGGTCTCAAGAAATCAAAAGACAGAGGTACCAAGCTCAA	1698
Dd	1618	GATATCAGAGACAGACCGGAGGTCTCAAGAAATCAAAAGACAGAGGTACCAAGCTCAA	1677
Qy	1699	GTGGAGAAATGAGGTTGATGATGATCAAGTGTGGAGAGGATCTTGTCTCAGCAAGAGA	1758
Dd	1678	GTGGAGAAATGAGGTTGATGATGATCAAGTGTGGAGAGGATCTTGTCTCAGCAAGAGA	1737
Qy	1759	CGAGTGTATCTCTCAAGAACTGAGCTGAGAGAGTCTCGGCTTGTGCTGAAGAAATTCAG	1818
Dd	1738	CGAGTGTATCTCTCAAGAACTGAGCTGAGAGAGTCTCGGCTTGTGCTGAAGAAATTCAG	1797
Qy	1819	CGGAAAGCCACAGAAATGTCAGCATAACTGTTGAAGGCTAAGATCAAGGAGGCTGAA	1878
Dd	1798	CGGAAAGCCACAGAAATGTCAGCATAACTGTTGAAGGCTAAGATCAAGGAGGCTGAA	1857
Qy	1879	GTGGGAGAAATGCGAAACTGGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1938
Dd	1858	GTGGGAGAAATGCGAAACTGGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1917
Qy	1939	CTTCCAGAGAAACTGGAGAGGCTG-----	1963
Dd	1918	CTTCCAGAGAAACTGGAGAGGCTGTAAAGCCAGCAGCGAGGCCACAGAGCTGTGCAG	1977
Qy	1964	-----	2010
Dd	1978	AAATATCCGCGCAGGCAAGAGGAGCGAGCGAGAGGAGCTGGAGAGCTGAGAAACCGAGAG	2037
Qy	2011	GATTCTTCTGAAGGCTACAGAAAGAGCTGGTGGAGGCTGAGGAAACCGCGCATTTCTCTG	2070
Dd	2038	GATTCTTCTGAAGGCTACAGAAAGAGCTGGTGGAGGCTGAGGAAACCGCGCATTTCTCTG	2097
Qy	2071	GAGAACAGGTAAGAGACTAGAGACCATGGAGCGGTAGAGAAACAGACTGAAGATGAC	2130
Dd	2098	GAGAACAGGTAAGAGACTAGAGACCATGGAGCGGTAGAGAAACAGACTGAAGATGAC	2157
Qy	2131	ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGGAGCTCGAAGAG	2190
Dd	2158	ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGGAGCTCGAAGAG	2217
Qy	2191	AAACATCGGGAGGCCCAAGTCTCAGCCGAGCCTTAGAAGTGCACCTGAAAACAGAAAGAG	2250
Dd	2218	AAACATCGGGAGGCCCAAGTCTCAGCCGAGCCTTAGAAGTGCACCTGAAAACAGAAAGAG	2277
Qy	2251	CAGCACTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAAGAAACACCTGGCTGAC	2310
Dd	2278	CAGCACTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAAGAAACACCTGGCTGAC	2337
Qy	2311	AAGGAGACACTGGAGAAACATGATGACAGACACGAGAGGAGGCCCATGAGAAGGGCAAA	2370
Dd	2338	AAGGAGACACTGGAGAAACATGATGACAGACACGAGAGGAGGCCCATGAGAAGGGCAAA	2397
Qy	2371	ATTCTCAGCGAACAGAGGGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAA	2430
Dd	2398	ATTCTCAGCGAACAGAGGGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAA	2457
Qy	2431	CAGAGGATTTGGAACTGTCTGAGGCCAATAAATTTGACGCAATATAGCAGTCTTTTACC	2490
Dd	2458	CAGAGGATTTGGAACTGTCTGAGGCCAATAAATTTGACGCAATATAGCAGTCTTTTACC	2517
Qy	2491	CAAAAGGAACATGAGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTAC	2550
Dd	2518	CAAAAGGAACATGAGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTAC	2577
Qy	2551	CTGAGAGACAGGCTGGAAAGTTGGAGGCCAGAACCGAAATCTGGAGGAGCAGCTGGAG	2610
Dd	2578	CTGAGAGACAGGCTGGAAAGTTGGAGGCCAGAACCGAAATCTGGAGGAGCAGCTGGAG	2637
Qy	2611	AAGATCAGGCCACCAAGACCAAGTGAACAAGATCGGCTGTGGAATCTGGAGACAAAGATTG	2670
Dd	2638	AAGATCAGGCCACCAAGACCAAGTGAACAAGATCGGCTGTGGAATCTGGAGACAAAGATTG	2697
Qy	2671	CGGAGAGTCACTTAGAGCAAGAGGACAGAACTGAGAGCTCAAGCCGAGCTCAGAGAG	2730
Dd	2698	CGGAGAGTCACTTAGAGCAAGAGGACAGAACTGAGAGCTCAAGCCGAGCTCAGAGAG	2757
Qy	2731	CTACAGCTCTCCCTGCAAGGAGCGGAGTCAAGTGTGACGCTGCAAGGCTGCAAGGCGG	2790
Dd	2758	CTACAGCTCTCCCTGCAAGGAGCGGAGTCAAGTGTGACGCTGCAAGGCTGCAAGGCGG	2817
Qy	2791	GCCCTGAGAGCCAGCTTCCGCAAGGCGGAGTCAAGTGTGACGCTGCAAGGAGAGCT	2850
Dd	2818	GCCCTGAGAGCCAGCTTCCGCAAGGCGGAGTCAAGTGTGACGCTGCAAGGAGAGCT	2877
Qy	2851	GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCGCAATTTGATGCT	2910
Dd	2878	GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCGCAATTTGATGCT	2937
Qy	2911	CTTCGTAAACAGCTTACTGTAAATCAGACAGCTTGAGGAGAGCTAAACAGCTGACCGAG	2970
Dd	2938	CTTCGTAAACAGCTTACTGTAAATCAGACAGCTTGAGGAGAGCTAAACAGCTGACCGAG	2997
Qy	2971	GACAAACGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT	3030
Dd	2998	GACAAACGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT	3057
Qy	3031	GGCCCAACAGCAGAGATTGTACAACTCGAAGTGAAGTGAACCTCTCCGCGGGAGATC	3090
Dd	3058	GGCCCAACAGCAGAGATTGTACAACTCGAAGTGAAGTGAACCTCTCTCCGCGGGAGATC	3117
Qy	3091	ACGAAACGAGAGATGCAAGCTTACAGCCAGAAAGCAACGATGGAGGCTTGAAGACCAAG	3150
Dd	3118	ACGAAACGAGAGATGCAAGCTTACAGCCAGAAAGCAACGATGGAGGCTTGAAGACCAAG	3177
Qy	3151	TGCACCATGCTGGAGGACAGGCTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTAGAA	3210
Dd	3178	TGCACCATGCTGGAGGACAGGCTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTAGAA	3237
Qy	3211	AAAGAGCGGAGTGGAGGCTTGAGGAGGCTTCTGGGTGATGAGAAATCCCAAGTTTGGAG	3270
Dd	3238	AAAGAGCGGAGTGGAGGCTTGAGGAGGCTTCTGGGTGATGAGAAATCCCAAGTTTGGAG	3297
Qy	3271	TGTCGGGTTTGAGAGCTGCAGAGATGCTGACACCGAGAAACAGAGCGGGCGAGAGCC	3330
Dd	3298	TGTCGGGTTTGAGAGCTGCAGAGATGCTGACACCGAGAAACAGAGCGGGCGAGAGCC	3357
Qy	3331	GATCAGCGGATCACCGAGTCTCCGAGGCTGGAGCTGGAGTGAAGGAGAGCAAGGCT	3390
Dd	3358	GATCAGCGGATCACCGAGTCTCCGAGGCTGGAGCTGGAGTGAAGGAGAGCAAGGCT	3417
Qy	3391	GAGATTCTCGCTCTGACAGGCTCTCAAGAGCAGAAAGTGAAGGCCGAGAGCTTCTCT	3450
Dd	3418	GAGATTCTCGCTCTGACAGGCTCTCAAGAGCAGAAAGTGAAGGCCGAGAGCTTCTCT	3477
Qy	3451	GACAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAATGAATGATGCTTGA	3510
Dd	3478	GACAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAATGAATGATGCTTGA	3537

QY 3511 CAGCAGAAAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAAGAGCAGCCAAA 3570
DB 3538 CAGCAGAAAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAAGAGCAGCCAAA 3597
QY 3571 TTACAGCAGCAGATGGACCTGCAGAAATAATCAATTTTCGCTCTGACTCAAGACTGCAA 3630
DB 3598 TTACAGCAGCAGATGGACCTGCAGAAATAATCAATTTTCGCTCTGACTCAAGACTGCAA 3657
QY 3631 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTATCAGCTG 3690
DB 3658 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTATCAGCTG 3717
QY 3691 GAAACCAATTCAGGTTCTCTATCTCATGAAAGAGTGAAGTGAAGGCACTATTTCTCAA 3750
DB 3718 GAAACCAATTCAGGTTCTCTATCTCATGAAAGAGTGAAGTGAAGGCACTATTTCTCAA 3777
QY 3751 CAAACCAAACTCATGATTTTCTGCAAGCCAAAATGACCAACTCTCTAAAAGAAAAAG 3810
DB 3778 CAAACCAAACTCATGATTTTCTGCAAGCCAAAATGACCAACTCTCTAAAAGAAAA -- 3835
QY 3811 GGTTTATTAGTCGACGGAAGAGGCCCTGCTTTTACCACACAGGTTCTCTGCAGTAC 3870
DB 3836 -----AGTGCTCTGCAGTAC 3852
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGACAGAGCTAGAGGAAGCC 3930
DB 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGACAGAGCTAGAGGAAGCC 3912
QY 3931 CTTGAGAAAGACCGCATGAGCTCGGTCGCGCCGGAGAGAGCTGCCACCGCAAGCA 3990
DB 3913 CTTGAGAAAGACCGCATGAGCTCGGTCGCGCCGGAGAGAGCTGCCACCGCAAGCA 3972
QY 3991 ACGGACCAACCAACCCATCCAGCCAGACCAACCGCAGGAGCAGATCGCATGTCGCC 4050
DB 3973 ACGGACCAACCAACCCATCCAGCCAGACCAACCGCAGGAGCAGATCGCATGTCGCC 4032
QY 4051 ATGTCGGTGGCCAGAGCACCAGCCAGTCGCATGAGGCTGTGCGCCCGCCATCCAGC 4110
DB 4033 ATGTCGGTGGCCAGAGCACCAGCCAGTCGCATGAGGCTGTGCGCCCGCCATCCAGC 4092
QY 4111 CGCAGAAAGAGCTTCACTCCAGAGGATTTAGTCGGCTCTTAAGNAGCCATGAC 4170
DB 4093 CGCAGAAAGAGCTTCACTCCAGAGGATTTAGTCGGCTCTTAAGNAGCCATGAC 4152
QY 4171 CACAATATTCCTCAGCGATTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGCTG 4230
DB 4153 CACAATATTCCTCAGCGATTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGCTG 4212
QY 4231 TGTCTGGATACCGTGACCTTTGAGCCGAGGATCCAAATGCTCGAATGTCAGGTGATG 4290
DB 4213 TGTCTGGATACCGTGACCTTTGAGCCGAGGATCCAAATGCTCGAATGTCAGGTGATG 4272
QY 4291 TGTCAACCCCAAGTGTCCAGCTCTTCCAGCCACCTGCGGCTTGCTGCTGAATATGCC 4350
DB 4273 TGTCAACCCCAAGTGTCCAGCTCTTCCAGCCACCTGCGGCTTGCTGCTGAATATGCC 4332
QY 4351 ACACATTCACCGAGGCTTCTCGCTGACAAATGAATCCAGGCTCCAGACCAAG 4410
DB 4333 ACACATTCACCGAGGCTTCTCGCTGACAAATGAATCCAGGCTCCAGACCAAG 4392
QY 4411 GAGCCAGCAGAGCTTGACCTTGGAAAGGTTGATGAAGTGCACGAGTAACAAACGA 4470
DB 4393 GAGCCAGCAGAGCTTGACCTTGGAAAGGTTGATGAAGTGCACGAGTAACAAACGA 4452
QY 4471 GGAACAAGGCTGGGACAGGAAGTACAATGTCCTGGAGGATCAAAAGTCTCTATTTAT 4530
DB 4453 GGAACAAGGCTGGGACAGGAAGTACAATGTCCTGGAGGATCAAAAGTCTCTATTTAT 4512
QY 4531 GACAATGAAGCAGAGAGCTTGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCC 4590
DB 4513 GACAATGAAGCAGAGAGCTTGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCC 4572
QY 4591 GACGGGATGTATCTATTATGTTGCTTCCGAACTCCGAAATACAGCCAAA 4650

DB 4573 GACGGGATGTATCTATTTCATGGTGGCTTGGTCTTCCGAACTCGAAATACAGCCAAA 4632
QY 4651 GCA ----- 4653
DB 4633 GCAGATGTCCTATACATACATGAAGATGAATCTCACCCGACACACCCTGCTGCTGCCGGG 4692
QY 4654 ----- 4653
DB 4693 AGAACCTCTTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCGCTGGTCAACGCTTA 4752
QY 4654 -----GAAAAAGCAGAAAGCTGATGCTAAACTG 4680
DB 4753 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGGAAAAAGCAGAAAGCTGATGCTAAACTG 4812
QY 4681 CTTGAAAACTCCCTGCTGAAAACTGGAAGGTGATGACCGTCTAGACATGAATGCTGACGCTG 4740
DB 4813 CTTGAAAACTCCCTGCTGAAAACTGGAAGGTGATGACCGTCTAGACATGAATGCTGACGCTG 4872
QY 4741 CCTTCAGTGACCAAGGTGCTGTTGTTGGGCAACCGAGGAAGGGCTCTACGCGCTGAATGTC 4800
DB 4873 CCTTCAGTGACCAAGGTGCTGTTGTTGGGCAACCGAGGAAGGGCTCTACGCGCTGAATGTC 4932
QY 4801 TTGAAAACTCCCTAAACCCATGTCCTAGGAATTCAGGAGCTTCCAAAATTTATTTATTC 4860
DB 4933 TTGAAAACTCCCTAAACCCATGTCCTAGGAATTCAGGAGCTTCCAAAATTTATTTATTC 4992
QY 4861 AAGGACCTGGAGAGCTTACTCATGATAGCAGGAAGAGCGGCACTGCTTTGTCGAC 4920
DB 4993 AAGGACCTGGAGAGCTTACTCATGATAGCAGGTGAAGCGGCACTGCTTTGTCGAC 5052
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGCACTCTCA 4980
DB 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGCACTCTCA 5112
QY 4981 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGCAGATGAGAAC 5040
DB 5113 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGCAGATGAGAAC 5172
QY 5041 GGGCTCTGCATCTCTGAGGCAATGCCAGCAAAAGTCTCATTTCCCTCAACGAAAC 5100
DB 5173 GGGCTCTGCATCTCTGAGGCAATGCCAGCAAAAGTCTCATTTCCCTCAACGAAAC 5232
QY 5101 CTCAGAAATATCTCATCTCGGAAAGATAGAGACCTCAGAGCCCTGCACTGTATCCAC 5160
DB 5233 CTCAGAAATATCTCATCTCGGAAAGATAGAGACCTCAGAGCCCTGCACTGTATCCAC 5292
QY 5161 TTCAACCAATACAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
DB 5293 TTCAACCAATACAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5352
QY 5221 TACAGCTCGAGGAATTTCTGGATAAGAAATGAOCATTTCTTTGGCAGCTGTGTTTGGC 5280
DB 5353 TACAGCTCGAGGAATTTCTGGATAAGAAATGAOCATTTCTTTGGCAGCTGTGTTTGGC 5412
QY 5281 GCCTTTCCAGACAGTTCCTGCTCAATCGTCAGGTGAACAGCGCAGGCGAGAG 5340
DB 5413 GCCTTTCCAGACAGTTCCTGCTCAATCGTCAGGTGAACAGCGGCGAGGCGAGAG 5472
QY 5341 GAGTACTTGTGTTTCCACGAATTTGGAGTGTTCGTGAATTTTACGGAAGACGTGAC 5400
DB 5473 GAGTACTTGTGTTTCCACGAATTTGGAGTGTTCGTGAATTTTACGGAAGACGTGAC 5532
QY 5401 CGCAGACAGATCTCAAGTGGAGTCCCTTACCTTTGGCTTCCCTACAGAGACCTAT 5460
DB 5533 CGCAGACAGATCTCAAGTGGAGTCCCTTACCTTTGGCTTCCCTACAGAGACCTAT 5592
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAAGTAAATGAGATCCAGGCACTCTCAGCA 5520
DB 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAAGTAAATGAGATCCAGGCACTCTCAGCA 5652
QY 5521 GGGACCCCTGCGAGGCTTACCTGGACATCCGAAACCCGCGCTTACCTGGGCTTCCCTT 5580

QY 2731 CTAAGCTCTCCCTGAGAGCGCGAGTCAAGTTGACAGCCCTCGAGCTGCACGGGG 2790
Db 2758 CTACAGCTCTCCCTGAGGAGCGCGAGTCAAGTTGACAGCCCTGAGGCTGCACGGGG 2817
QY 2791 GCCCTGGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACCAACAGAGCT 2850
Db 2818 GCCCTGGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACCAACAGAGCT 2877
QY 2851 GAAGAGAGATCCAGCACTCAACGACATAGAGATGAATCAGGCGAAATTTGATGCT 2910
Db 2878 GAAGAGAGATCCAGCACTCAACGACATAGAGATGAATCAGGCGAAATTTGATGCT 2937
QY 2911 CTTCTGTAAGAGCTGTACTGTATACAGACCTGGAGAGAGCTAAACAGCTGACCGAG 2970
Db 2938 CTTCTGTAAGAGCTGTACTGTATACAGACCTGGAGAGAGCTAAACAGCTGACCGAG 2997
QY 2971 GACAAGCTGAACTCAACCAACAAACTTCTACTTGTCCAAACCACTCGATGAGCTTCT 3030
Db 2998 GACAAGCTGAACTCAACCAACAAACTTCTACTTGTCCAAACCACTCGATGAGCTTCT 3057
QY 3031 GGCGCCAAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3090
Db 3058 GGCGCCAAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3117
QY 3091 ACGERAACGAGAGATGAGCTTACCGCCAGAGCAAAACGATGAGGCTCTGAGAGCCAG 3150
Db 3118 ACGERAACGAGAGATGAGCTTACCGCCAGAGCAAAACGATGAGGCTCTGAGAGCCAG 3177
QY 3151 TGCAACCATGCTGGAGAAACAGGTCATGAGTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210
Db 3178 TGCAACCATGCTGGAGAAACAGGTCATGAGTTGGAGGCCCTAAACGATGAGCTGTAGAA 3237
QY 3211 AAGAGCGGAGTGGAGGCGCTGGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGA 3270
Db 3238 AAGAGCGGAGTGGAGGCGCTGGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGA 3297
QY 3271 TGTCCGGTTCCGAGAGTGCAGAGAAATGCTGGACACCGAGAAACAGAGCGGAGAGCC 3330
Db 3298 TGTCCGGTTCCGAGAGTGCAGAGAGTGTGGACACCGAGAAACAGAGCGGAGAGCC 3357
QY 3331 GATCAGCGATCACCAGTCTCCAGAGTGTGGAGCTGGAGTGGAGAGACAGAGCT 3390
Db 3358 GATCAGCGATCACCAGTCTCCAGAGTGTGGAGCTGGAGTGGAGAGACAGAGCT 3417
QY 3391 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3450
Db 3418 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3477
QY 3451 GACAAGCTCAATGACCTGGAGAGAAAGCATGCTTGAATGAATGCCGAAAGCTTA 3510
Db 3478 GACAAGCTCAATGACCTGGAGAGAAAGCATGCTTGAATGAATGCCGAAAGCTTA 3537
QY 3511 CAGCAGAGCTGAGAGTGAACAGAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAA 3570
Db 3538 CAGCAGAGCTGAGAGTGAACAGAGGCTCAACAGAGGCTTCTGGAGAGCAAGCCAA 3597
QY 3571 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTCCGCTCTGACTCAAGGAGTGA 3630
Db 3598 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTCCGCTCTGACTCAAGGAGTGA 3657
QY 3631 GAAGCTCTAGATCGGCTGTATCTATGAGAGACAGAAAGAGTGAATTGAGTATCAGCTG 3690
Db 3658 GAAGCTCTAGATCGGCTGTATCTATGAGAGACAGAAAGAGTGAATTGAGTATCAGCTG 3717
QY 3691 GAAAAATTCAGGTTCTCTATTCTCATGAAAGGTGAAAAATGAAAGGCACTATTCTCAA 3750
Db 3718 GAAAAATTCAGGTTCTCTATTCTCATGAAAGGTGAAAAATGAAAGGCACTATTCTCAA 3777
QY 3751 CAAACCAAACTCATGATTCTTTCGAAAGCAAAATGACCAACCTGCTTAAAGAAAAAG 3810
Db 3778 CAAACCAAACTCATGATTCTTTCGAAAGCAAAATGACCAACCTGCTTAAAGAAAAAG - - 3835

QY 3811 GGTTTATTTAGTCGACGAAAGAGGACCTCTGCTTTACCCACACAGGTTCTCTGAGTAC 3870
Db 3836 -----AGGTGCTCTGAGTAC 3852
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCC 3930
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCC 3912
QY 3931 CTTTCAAGAGACCCCATGAGCTCCGCTCGCCCGGAGGAAGCTGCCCAACGAAAGCA 3990
Db 3913 CTTTCAAGAGACCCCATGAGCTCCGCTCGCCCGGAGGAAGCTGCCCAACGAAAGCA 3972
QY 3991 ACAGACACACACACACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATCTCCGCG 4050
Db 3973 ACAGACACACACACACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATCTCTGCG 4032
QY 4051 ATCTGTGCGGTGCGCCAGAGACACCCAGTGCCTTGAAGCTGCTGGCCCGCCCATCCAGC 4110
Db 4033 ATCTGTGCGGTGCGCCAGAGACACCCAGTGCCTTGAAGCTGCTGGCCCGCCCATCCAGC 4092
QY 4111 CGCAGAAAGAGAGTCTTCAACTCCAGAGAAATTTAGTGGGCTCTTAAGGAACGATGAC 4170
Db 4093 CGCAGAAAGAGAGTCTTCAACTCCAGAGAAATTTAGTGGGCTCTTAAGGAACGATGAC 4152
QY 4171 CACAATATTCTCACCAGTCAAGCTAGGACTGAACATGCGAGCAACAAAGTGTGCTGTG 4230
Db 4153 CACAATATTCTCACCAGTCAAGCTAGGACTGAACATGCGAGCAACAAAGTGTGCTGTG 4212
QY 4231 TGTCTGAGTACCGTGCACTTTGGACGCGCAGCATCCAAATGTCTGCAATGTCTGAGTATG 4290
Db 4213 TGTCTGAGTACCGTGCACTTTGGACGCGCAGCATCCAAATGTCTGCAATGTCTGAGTATG 4272
QY 4291 TGTCAACCCAGTGTCTCCAGTGTCTCCAGCCAGCTTGGGCTTGGCTGTGATATGCTG 4350
Db 4273 TGTCAACCCAGTGTCTCCAGTGTCTCCAGCCAGCTTGGGCTTGGCTGTGATATGCTG 4332
QY 4351 ACACATTTACCGAGGCTTCTGCGCTGACAAAAATGAACCTCCAGAGTCTCCAGACCAAG 4410
Db 4333 ACACATTTACCGAGGCTTCTGCGCTGACAAAAATGAACCTCCAGAGTCTCCAGACCAAG 4392
QY 4411 GAGCCAGCAGCAGCTTGCACCTGGAGAGGTGAGTGAAGTGGCCAGGAATAACAAAGCA 4470
Db 4393 GAGCCAGCAGCAGCTTGCACCTGGAGAGGTGAGTGAAGTGGCCAGGAATAACAAAGCA 4452
QY 4471 GGAAGCAAGGCTGGGACAGGAAGTCAATTTCTCTGAGGAGTCAAAAGTCTCTATTAT 4530
Db 4453 GGAAGCAAGGCTGGGACAGGAAGTCAATTTCTCTGAGGAGTCAAAAGTCTCTATTAT 4512
QY 4531 GACAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4590
Db 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4572
QY 4591 GACGGGATGTATCTATTCTATGCTGCTTGGCTTCCGAACTCGCAATACAGCCAAA 4650
Db 4573 GACGGGATGTATCTATTCTATGCTGCTTGGCTTCCGAACTCGCAATACAGCCAAA 4632
QY 4651 GCA----- 4653
Db 4633 GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGACACCACTCTGCTGCGCCGG 4692
QY 4654 ----- 4653
Db 4693 AGAACCCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGAGCTGGGTACCGCCCTTA 4752
QY 4654 -----GAAAAAGCAGAAAGTGAATGCTGCTTAAACTG 4690
Db 4753 GAATCAGTTGTGCGAGTGGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTTAACTG 4812
QY 4681 CTTGGAACTCCCTGTGTAACCTGGAAGGTGATGACCTGTAGACATGAATGCAAGCTG 4740
Db 4813 CTTGGAACTCCCTGTGTAACCTGGAAGGTGATGACCTGTAGACATGAATGCAAGCTG 4872
QY 4741 CCCTTCAGTGACAGGTTGGTGTGGTGGGACCCGAGGAGGGCTCTACGCTTGAATGTC 4800

[illegible]

```

RESULT 10
US-09-964-956-8
; Sequence 8, Application US/09964956
; Publication No. US20040043926A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No US20040043926A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964,956
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276,667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294,823
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304,868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6201
; TYPE: DNA

```

; ORGANISM: Homo sapiens			
US-09-964-956-8			
Query Match 85.6%; Score 5629.8; DB 13; Length 6201;			
Best Local Similarity 95.7%; Pred. No. 0;			
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;			
QY	19	ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGGTGTCTGCTGAACCCATT	78
DB	1	ATGTTGAAGTTCNAATATGAGCGCGGAATCCCTTTGGATGCTGGTGTCTGCTGAACCCATT	60
QY	79	GCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA	138
DB	61	GCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA	120
QY	139	CAGCAGATGTCCTCTCTTCCGAGAGGATATATAGATGCCCTCTTTCTCTTTGAA	198
DB	121	CAGCAGATGTCCTCTCTTCCGAGAGGATATATAGATGCCCTCTTTCTCTTTGAA	180
QY	199	GAATGAGTCAAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGGAATAT	258
DB	181	GAATGAGTCAAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGGAATAT	240
QY	259	TCCGACACCATAGCTAGTACAGAGCTCCAGCCTTCGGCAAGACTTCGAAGTCAGA	318
DB	241	TCCGACACCATAGCTAGTACAGAGCTCCAGCCTTCGGCAAGACTTCGAAGTCAGA	300
QY	319	AGTCTTTGAGTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAGAAACCAACCGG	378
DB	301	AGTCTTTGAGTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAGAAACCAACCGG	360
QY	379	GACATCTATGCTATGAAGATGAAGAAGAGGCTTTATGCGCCAGAGAGAGGTTTCA	438
DB	361	GACATCTATGCTATGAAGATGAAGAAGAGGCTTTATGCGCCAGAGAGAGGTTTCA	420
QY	439	TTTTTTGAGAGAGGCGGAACATATATCTCGAAGCACAAAGCCGTTGATCCCAATT	498
DB	421	TTTTTTGAGAGAGGCGGAACATATATCTCGAAGCACAAAGCCGTTGATCCCAATT	480
QY	499	CAGTATGCTTTTCAAGCAAAAATACCTTTATCTGATGAGGAATATCAGCCTGGAGG	558
DB	481	CAGTATGCTTTTCAAGCAAAAATACCTTTATCTGATGAGGAATATCAGCCTGGAGG	540
QY	559	GACTTGTCTCACTTTTGAATAGATAGAGGACCAAGTAGATGAAGAACTGTATGAGTTT	618
DB	541	GACTTGTCTCACTTTTGAATAGATAGAGGACCAAGTAGATGAAGAACTGTATGAGTTT	600
QY	619	TACCTAGCTGAGTGAATTTGGCTGTTTACAGCGTTTCATCTGATGGATACGTGCATCGA	678
DB	601	TACCTAGCTGAGTGAATTTGGCTGTTTACAGCGTTTCATCTGATGGATACGTGCATCGG	660
QY	679	GACATCAAGCCTGAGAACATCTCGTTGACCGCAGGACACATCAAGCTGTGATTTT	738
DB	661	GACATCAAGCCTGAGAACATCTCGTTGACCGCAGGACACATCAAGCTGTGATTTT	720
QY	739	GGATCTGCGCGGAAATGAATCAAAACAGATGGTGAATGCCAACTCCCGATTGGGACC	798
DB	721	GGATCTGCGCGGAAATGAATCAAAACAGATGGTGAATGCCAACTCCCGATTGGGACC	777
QY	799	CCAGATTAATGCTCCTGAACTGTGACTGTGATGAACGGGATGGAAGGACACCTAC	858
DB	778	CCAGATTAATGCTCCTGAACTGTGACTGTGATGAACGGGATGGAAGGACACCTAC	837
QY	859	GGCTCGGACTGTGACTGGTGGTCACTGGGCGGTGATTCCTATGAGATGATTTATGGGAGA	918
DB	838	GGCTCGGACTGTGACTGGTGGTCACTGGGCGGTGATTCCTATGAGATGATTTATGGGAGA	897
QY	919	TCCCTCTCCAGAGGGAACCTCTGCCAGAACCTTTCAATAACATATGAATTTCCAGCGG	978
DB	898	TCCCTCTCCAGAGGGAACCTCTGCCAGAACCTTTCAATAACATATGAATTTCCAGCGG	957
QY	979	TTTTTGAATTTCCAGATGACCCCAAGTGACGAGTACTTTCTTGATCTGATTCATAAGC	1038
DB			

958	TTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTCTTCTGATCTGATTCACAAAGC	1017
1039	TTGTTGTCGCGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCTGCTATCTTCTTC	1098
1018	TTGTTGTCGCGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCTGCTATCTTCTTC	1077
1099	TCTAAATTTGACTGGAACCAACATTCGTAACCTCTCTCCCTTCTGTTCCCACTCAAG	1158
1078	TCTAAATTTGACTGGAACCAACATTCGTAACCTCTCTCCCTTCTGTTCCCACTCAAG	1137
1159	TCTGACGATGACACTCCTCAATTTTGTATGAACAGAGAGAAATTCGTGGGTTCATCTCT	1218
1138	TCTGACGATGACACTCCTCAATTTTGTATGAACAGAGAGAAATTCGTGGGTTCATCTCT	1197
1219	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTCAAGAACTGCCGTTTGTGGGTTTTTCG	1278
1198	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTCAAGAACTGCCGTTTGTGGGTTTTTCG	1257
1279	TACAGCAAGGCACTGGGGATTCCTTGGTAGATCTGAGTCTGTGTGTCGGGTCTGAGCTCC	1338
1258	TACAGCAAGGCACTGGGGATTCCTTGGTAGATCTGAGTCTGTGTGTCGGGTCTGAGCTCC	1317
1339	CCTGCCAGACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTTACAGAGC	1398
1318	CCTGCCAGACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTTACAGAGC	1377
1399	TCTCAGGACAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTTACATCGAGAGTCTCA	1458
1378	TCTCAGGACAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTTACATCGAGAGTCTCA	1437
1459	GAGTGGAGGCTGTCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1518
1438	GAGTGGAGGCTGTCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1497
1519	TCCCTCTCGAGCAGGACCTTGCTTACCTCATCATCAGAAATGCAGTAGCTTAAAGCGAAGT	1578
1498	TCCCTCTCGAGCAGGACCTTGCTTACCTCATCATCAGAAATGCAGTAGCTTAAAGCGAAGT	1557
1579	TTGAGCAAGCAGGATGGAGGTGTCACAGAGGATCAAAAGCACTGCGAGCTTCTCCAT	1638
1558	TTGAGCAAGCAGGATGGAGGTGTCACAGAGGATCAAAAGCACTGCGAGCTTCTCCAT	1617
1639	GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAAGCTCAA	1698
1618	GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAAGCTCAA	1677
1699	GTGCAAGAAATGAGTTGATGAATCACTTGGAGAGGATCTTGTCTCAGCAAGAAAGA	1758
1678	GTGCAAGAAATGAGTTGATGAATCACTTGGAGAGGATCTTGTCTCAGCAAGAAAGA	1737
1759	CGGAGTATCTCTACGAACTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGAAATTCAG	1818
1738	CGGAGTATCTCTACGAACTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGAAATTCAG	1797
1819	CGGAAACCGACAGAAATGTGAGCATAACTGTTGAAGCTAAGGATCAAGGAAAGCTGAA	1878
1798	CGGAAACCGACAGAAATGTGAGCATAACTGTTGAAGCTAAGGATCAAGGAAAGCTGAA	1857
1879	GTGGGAGAAATATCGAAACTTGGAGAGATCAATGCTGAGCAGGAGCTCAAAATTCAGGAG	1938
1858	GTGGGAGAAATATCGAAACTTGGAGAGATCAATGCTGAGCAGGAGCTCAAAATTCAGGAG	1917
1939	CTCCACAGAGAACTGGAGAAAGGCT-----	1962
1918	CTCCACAGAGAACTGGAGAAAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGTGCGAG	1977
1963	-----GCAAGGAGCGAGCGAGGAGCTGGAGAGCTGGAGAGCTGCAGAACCGAGAG	2010
1978	AATATCCCGCAGGCAAGGAGCGAGCGAGGAGCTGGAGAGCTGCAGAACCGAGAG	2037
2011	GATTCTTTCTGAAGGCTCAGAAAGAGAGTCTGTTGGAAGCTCAGGAAACGCGCGCTTCTCTG	2070
2038	GATTCTTTCTGAAGGCTCAGAAAGAGAGTCTGTTGGAAGCTGAGGAACGCGCGCTTCTCTG	2097

Db 4213 TGCTGGATACCGTGACCTTTGGAGCCAGGCAATCCAAATGCTAGATGTGAGGTGATG 4272
QY 4291 TGTACCCCAAGTGTCCAGTGTTCGCCAGCCACCTGCGGTGCTGCTGTAATATGCC 4350
Db 4273 TGTACCCCAAGTGTCCAGTGTTCGCCAGCCACCTGCGGTGCTGCTGTAATATGCC 4332
QY 4351 ACACACTTCACCGAGGCTTCCGCTGACAAAATGAATCCCCAGGTCTCCAGACCAAG 4410
Db 4333 ACACACTTCACCGAGGCTTTCGCCGTGACAAAATGAATCCCCAGGTCTCCAGACCAAG 4392
QY 4411 GAGCCAGCAGCAGCTTGACCTTGGAGGGTGGATGAAGGTGCCAGGAATTAACAAAGA 4470
Db 4393 GAGCCAGCAGCAGCTTGACCTTGGAGGGTGGATGAAGGTGCCAGGAATTAACAAAGA 4452
QY 4471 GGNACAGAGGCTGGACAGGAAGTACATTTGCTGGAGGGATCAAAAGTCTCATTTAT 4530
Db 4453 GGNACAGAGGCTGGACAGGAAGTACATTTGCTGGAGGGATCAAAAGTCTCATTTAT 4512
QY 4531 GACAAATGAAGCAGAGAACTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCTTCCC 4590
Db 4513 GACAAATGAAGCAGAGAACTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCTTCCC 4572
QY 4591 GACGGGGATGTATCTATTATCATGGTGCCTGTGGTCTTCCGAACTCGCAAAATACGCCAAA 4650
Db 4573 GACGGGGATGTATCTATTATCATGGTGCCTGTGGTCTTCCGAACTCGCAAAATACGCCAAA 4632
QY 4651 GCA----- 4653
Db 4633 GCAGATGTCCTCATCATCTGAAGATGGAATCTCACCCGCCACACCACTGCTGGCGCCGGG 4692
QY 4654 ----- 4653
Db 4693 AGAACCTCTACTTGTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCCCTTA 4752
QY 4654 ----- 4653
Db 4753 GAATCAGTTGTCCGAGTGGAGAGTTTCTACGGGAAAGAGAGAGCTGATGCTAAACTG 4812
QY 4681 CTTTGGAACTCCCTGCTGAAATCTGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4740
Db 4813 CTTTGGAACTCCCTGCTGAAATCTGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4872
QY 4741 CCCTTCAGTGACAGTGTGTGTGGGCAACGAGAAAGGCTCTACGCCCTGAATGTC 4800
Db 4873 CCCTTCAGTGACAGTGTGTGTGGGCAACGAGAAAGGCTCTACGCCCTGAATGTC 4932
QY 4801 TTGAAAACTCCCTAACCCATGTCACAGAAATTGGAGCAGTCTTCCAAATTTATATTATC 4860
Db 4933 TTGAAAACTCCCTAACCCATGTCACAGAAATTGGAGCAGTCTTCCAAATTTATATTATC 4992
QY 4861 AAGGACCTGGAGAACTACTCATGATAGCAGGAGAGAGCGGCGACTGTGTCTTGTGGAC 4920
Db 4993 AAGGACCTGGAGAACTACTCATGATAGCAGGAGAGAGCGGCGACTGTGTCTTGTGGAC 5052
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCGACCTGCTGCCAGCCGACATCTCA 4980
Db 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCGACCTGCTGCCAGCCGACATCTCA 5112
QY 4981 CCCAACTTTTGAAGCTGTCAAGGCTGCCACTTGTGTGGGCGAGCAAGATTGAGAAC 5040
Db 5113 CCCAACTTTTGAAGCTGTCAAGGCTGCCACTTGTGTGGGCGAGCAAGATTGAGAAC 5172
QY 5041 GGGCTCTGATCTGTGACGCCATGCCAGCAAGTGTGCTATTCCTCGGTACAAAGCAAAAC 5100
Db 5173 GGGCTCTGATCTGTGACGCCATGCCAGCAAGTGTGCTATTCCTCGGTACAAAGCAAAAC 5232
QY 5101 CTCAGCAATACTGCATCCGGAAGAGATAGAGACTCAGAGCCCTGAGAGTGTATCCAC 5160
Db 5233 CTCAGCAATACTGCATCCGGAAGAGATAGAGACTCAGAGCCCTGAGAGTGTATCCAC 5292
QY 5161 TTCACCAATTACAGTATCTCTATTGGAACCAATAAAATTTACGAAATGCATGAGAGAG 5220

Db 5293 TTCACCAATTACAGTATCCTCATTTGGAACCAATAAAATTTACGAAATCGCATGAGCAG 5352
QY 5221 TACACGCTCGAGGAATTTCTGGATAAGAAATGAACATTTCTTGGCACCTGTGTGTTGCC 5280
Db 5353 TACACGCTCGAGGAATTTCTGGATAAGAAATGAACATTTCTTGGCACCTGTGTGTTGCC 5412
QY 5281 GCCTTTCCCAACAGCTTCCCTGTCTCAATCGTGAGTGAACGCGGAGGCGACGGAGAG 5340
Db 5413 GCCTTTCCCAACAGCTTCCCTGTCTCAATCGTGAGTGAACGCGGAGGCGACGGAGAG 5472
QY 5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTTACGGAAGAAGCTAGC 5400
Db 5473 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTTACGGAAGAAGCTAGC 5532
QY 5401 CGCACAGACGATCTCAAGTGGAGTTCGTTTACCTTTGGCCCTTTGCTACAGAGAACCTTAT 5460
Db 5533 CGCACAGACGATCTCAAGTGGAGTTCGTTTACCTTTGGCCCTTTGCTACAGAGAACCTTAT 5592
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGACACGCTCCTCAGCA 5520
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGACACGCTCCTCAGCA 5652
QY 5521 GGGACCTCTGCCCGAGCGTACCTGGACATCCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5580
Db 5653 GGGACCTCTGCCCGAGCGTACCTGGACATCCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5712
QY 5581 TCCTCAGAGAGATTTACTTTGGCTCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5640
Db 5713 TCCTCAGAGAGATTTACTTTGGCTCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5772
QY 5641 AAGGAAACCTTGTGAAGGATCGGCACTGAACACACCCGGGCGCTCCACCTCCCGC 5700
Db 5773 AAGGAAACCTTGTGAAGGATCGGCACTGAACACACCCGGGCGCTCCACCTCCCGC 5832
QY 5701 AGCAGCCCCACAGAGCGAGGCCACCCACCTGACAGAGACACATCACCAAGCGGTGGCC 5760
Db 5833 AGCAGCCCCACAGAGCGAGGCCACCCACCTGACAGAGACACATCACCAAGCGGTGGCC 5892
QY 5761 TCCAGCCAGCGCCCGCCGAGGCCCCAGACACCCGCGAGAGCCAAAGCACACCCACCGC 5820
Db 5893 TCCAGCCAGCGCCCGCCGAGGCCCCAGACACCCGCGAGAGCCAAAGCACACCCACCGC 5952
QY 5821 TACCGCAGGGGGGAGCGAGCTGCGAGGGAAGTCTCTGCGCCGCCCTGAGGCGGA 5880
Db 5953 TACCGCAGGGGGGAGCGAGCTGCGAGGGAAGTCTCTGCGCCGCCCTGAGGCGGA 6012
QY 5881 GAGAAGTCCCGCCCGGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 5940
Db 6013 GAGAAGTCCCGCCCGGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 6072
QY 5941 GAAGACAGCAGCGGGCGCGCTGCTGCGGGAGCCCTGAGGACCCCGCTGTCCAGGTG 6000
Db 6073 GAAGACAGCAGCGGGCGCGCTGCTGCGGGAGCCCTGAGGACCCCGCTGTCCAGGTG 6132
QY 6001 AACAGGGAAGAGGCA 6017
Db 6133 AACAGGTGAGGCAGCA 6149

RESULT 11
US-10-028-946-3
; Sequence 3, Application US/10028946
; Publication No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Fiddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27

1981	DB	AAATATCCCGCCAGGCAAGGAGCGAGCGAGGAGGAGCTGGAGAAAGCTGCAGAAACCGAGAG	2041
2011	QY	GAATTCCTCTGAGGCGATCAGAAAAGAAAGCTGCTGGAAGCTGAGGAACCGCGCATTCCTCTG	2070
2041	DB	GAATTCCTCTGAGGCGATCAGAAAAGAGAGCTGCTGGAGAGCTGAGGAACCGCGCCATTCCTCTG	2100
2071	QY	GAGAACAAAGGTTAAAGAGACTAGAGACCAATGGAGCGTAGAGAAAAACAGACTGAAGGATGAC	2130
2101	DB	GAGAACAAAGGTTAAAGAGACTAGAGACCAATGGAGCGTAGAGAAAAACAGACTGAAGGATGAC	2160
2131	QY	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAAGAG	2190
2161	DB	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAAGAG	2220
2191	QY	AAAATATCGGAGGCGCCAAAGTCTCAGCCGACGACCTAGAGTGCACCTGAAACAGAAAAGAG	2250
2221	DB	AAAATATCGGAGGCGCCAAAGTCTCAGCCGACGACCTAGAGTGCACCTGAAACAGAAAAGAG	2280
2251	QY	CAGCACATATGAGGAAAGAGATTAAAGTGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2310
2281	DB	CAGCACTATGAGGAAAGAGATTAAAGTGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2340
2311	QY	AAGGAGACACTGGAGAACATGATCGAGACAACAGAGGAGGAGGCCCATGAGAAAGGCGAAA	2370
2341	DB	AAGGAGACACTGGAGAACATGATCGAGACAACAGAGGAGGAGGCCCATGAGAAAGGCGAAA	2400
2371	QY	ATTCTCAGCGAACAGAGGCGCATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGGAA	2430
2401	DB	ATTCTCAGCGAACAGAGGCGCATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGGAA	2460
2431	QY	CAGAGGATTTGTGGAACTGTCTGAAAGCCAAATAAACTTCAGCAAAATAGCAGTCTTTTATCC	2490
2461	DB	CAGAGGATTTGTGGAACTGTCTGAAAGCCAAATAAACTTCAGCAAAATAGCAGTCTTTTATCC	2520
2491	QY	CAAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTATC	2550
2521	DB	CAAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTATC	2580
2551	QY	CTGGAGACACAGGCTGGGAACTGTGGAGCCGAGAAACCGAAAATCTGGAGGAGCAGCTGGAG	2610
2581	DB	CTGGAGACACAGGCTGGGAACTGTGGAGCCGAGAAACCGAAAATCTGGAGGAGCAGCTGGAG	2640
2611	QY	AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGTGGAACTGGAGACAAGATTG	2670
2641	DB	AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGTGGAACTGGAGACAAGATTG	2700
2671	QY	CGGAGGTTTCAGTCTAGAGCAGCGAGGAGCAAACTGGAGCTCAGCGCCAGCTCACAAGAG	2730
2701	DB	CGGAGGTTTCAGTCTAGAGCAGCGAGGAGCAAACTGGAGCTCAGCGCCAGCTCACAAGAG	2760
2731	QY	CTACAGCTCTCCCTCGAGGAGCGGAGTCAACAGTTGACAGCCCTGCAAGGCTGCACGGCGG	2790
2761	DB	CTACAGCTCTCCCTCGAGGAGCGGAGTCAACAGTTGACAGCCCTGCAAGGCTGCACGGCGG	2820
2791	QY	GCCCTGGAGAGCCAGCTTCGCCAGGCGAGACAGAGCTGGAAAGAGACCAAGAGAGCT	2850
2821	DB	GCCCTGGAGAGCCAGCTTCGCCAGGCGAGACAGAGCTGGAAAGAGACCAAGAGAGCT	2880
2851	QY	GAAGAGGAGATTCAGGCACTCACGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2910
2881	DB	GAAGAGGAGATTCAGGCACTCACGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2940
2911	QY	CTTCGTAAACAGCTGACTGTAAATCACAAGACTGGAGGAGCAGCTTAAACAGCTGACCGAG	2970
2941	DB	CTTCGTAAACAGCTGACTGTAAATCACAAGACTGGAGGAGCAGCTTAAACAGCTGACCGAG	3000
2971	QY	GACAAACGCTGAACCTCAACCAACCAAACTTCTACTTTGTCAAAACAACTCGATGAGGCTTCT	3030
3001	DB	GACAAACGCTGAACCTCAACCAACCAAACTTCTACTTTGTCAAAACAACTCGATGAGGCTTCT	3060
3031	QY	GGCGCCAAACGACGAGATTGTAACATCGCAAGTGAAGTGAACCATCTCCGCCGGGAGATC	3090
3061	DB	GGCGCCAAACGACGAGATTGTAACATCGCAAGTGAAGTGAACCATCTCCGCCGGGAGATC	3120

QY	3091	ACGGAACGAGAGATGCGACTTACAGCCAGAAGCAAAACGATGGAGGCTCTGAAGACCACG	3115
DB	3121	ACGGAACGAGAGATGCGACTTACAGCCAGAAGCAAAACGATGGAGGCTCTGAAGACCACG	3180
QY	3151	TGCACCATGCTGGAGGACAGGTCAATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA	3210
DB	3181	TGCACCATGCTGGAGGACAGGTCAATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA	3240
QY	3211	AAAGAGCGGCAGTGGAGGCGCTGGAGAGCGTCTCTGGGTGATGAGAAATCCCAAGTTTGAG	3270
DB	3241	AAAGAGCGGCAGTGGAGGCGCTGGAGAGCGTCTCTGGGTGATGAGAAATCCCAAGTTTGAG	3300
QY	3271	TGTCGGGTTCCGAGAGCTGCAGAGAAATCTCGACACCGAGAGAAACAGACAGGGCGAGAGCC	3330
DB	3301	TGTCGGGTTCCGAGAGCTGCAGAGAAATCTCGACACCGAGAGAAACAGACAGGGCGAGAGCC	3360
QY	3331	GATCAGCGGATCACCGAGTCTCGCCAGCTGGTGGAGCTGCGAGTGAAGGAGACAAAGGCT	3390
DB	3361	GATCAGCGGATCACCGAGTCTCGCCAGCTGGTGGAGCTGCGAGTGAAGGAGACAAAGGCT	3420
QY	3391	GAGATTCTCGCTTGCAGAGCTCTCAAAAGACAGAGCTGAAGCCGAGAGCCTCTCT	3450
DB	3421	GAGATTCTCGCTTGCAGAGCCTCTCAAAAGACAGAGCTGAAGCCGAGAGCCTCTCT	3480
QY	3451	GACAAGCTCAATGACCTGGAGAGAAACATGCTATGCTTGAATGAATGCCGAAGCTTA	3510
DB	3481	GACAAGCTCAATGACCTGGAGAGAAACATGCTATGCTTGAATGAATGCCGAAGCTTA	3540
QY	3511	CAGCAGAAGCTGGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGACGAAGCCAAA	3570
DB	3541	CAGCAGAAGCTGGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGACGAAGCCAAA	3600
QY	3571	TTACAGCAGCAGATGGAGCTGCGAAGAAATCACAATTTCCGCTGCTCAAGAGACTGCAG	3630
DB	3601	TTACAGCAGCAGATGGAGCTGCGAAGAAATCACAATTTCCGCTGCTCAAGAGACTGCAG	3660
QY	3631	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTAGTCACTTGGAGTATCAGCTG	3690
DB	3661	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTAGTCACTTGGAGTATCAGCTG	3720
QY	3691	GAAGACATTGAGTTCTCTATTCTCATGAAAAAGTGAAAAATGGAAGGCATTTCTCAA	3750
DB	3721	GAAGACATTGAGTTCTCTATTCTCATGAAAAAGTGAAAAATGGAAGGCATTTCTCAA	3780
QY	3751	CAAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGGACCAACTGCTAAAAAGAAAAAG	3810
DB	3781	CAAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGGACCAACTGCTAAAAAGAAAAAG	3838
QY	3811	GGTTTATTTAGTCGACGGAAAGAGACCCCTGCTTTACCCACACAGGTTCTCTGCGAGTAC	3870
DB	3839	-----AGGTTCTCTGCGAGTAC	3855
QY	3871	AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCACAGCTAGAGGAAGCC	3930
DB	3856	AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCACAGCTAGAGGAAGCC	3915
QY	3931	CTTCAGAAGACCCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCTGCCACCGCAAGCA	3990
DB	3916	CTTCAGAAGACCCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCTGCCACCGCAAGCA	3975
QY	3991	ACGGACCAACCCATCCAGCCAGCCACCGCAGGACGACAGATCGCCATGTCGGCC	4050
DB	3976	ACGGACCAACCCATCCAGCCAGCCACCGCAGGACGACAGATCGCCATGTCGGCC	4035
QY	4051	ATCGTGGGTGCGCAGAGCACCGCCAGTGCATGAGCCTGCTGGCCCGGCCATCCAGC	4110
DB	4036	ATCGTGGGTGCGCAGAGCACCGCCAGTGCATGAGCCTGCTGGCCCGGCCATCCAGC	4095
QY	4111	CGCAGAAAGAGTCTTCAATCCAGAGGAAATTTAGTCGGGCTCTTAAGGAAACGATGCAAC	4170
DB	4096	CGCAGAAAGAGTCTTCAATCCAGAGGAAATTTAGTCGGGCTCTTAAAGAAACGATGCAAC	4155

Qy 4171 CACAAATATTCCTCACCGATTCAACGTAGGACTGAACATCGGAGCCACAAAGTGTGCTGTG 4230
Db 4156 CACAAATATTCCTCACCGATTCAACGTAGGACTGAACATCGGAGCCACAAAGTGTGCTGTG 4215
Qy 4231 TGTCTGGATACCGTGGACATTTGGACCGCAGGCAATCCAAATGCTCGAATGTCCAGGTGATG 4290
Db 4216 TGTCTGGATACCGTGGACATTTGGACCGCAGGCAATCCAAATGCTCGAATGTCCAGGTGATG 4275
Qy 4291 TGTACCCCCAAGTGCTCCAGTCTTCCAGGACACCTGCGGCTTGCCTGCTGAATATGCC 4350
Db 4276 TGTACCCCCAAGTGCTCCAGTCTTCCAGGACACCTGCGGCTTGCCTGCTGAATATGCC 4335
Qy 4351 ACACACTTACCGAGGCTTCTCCGCTGACAAATGAATCTCCAGGCTCTCCAGACCAAG 4410
Db 4336 ACACACTTACCGAGGCTTCTCCGCTGACAAATGAATCTCCAGGCTCTCCAGACCAAG 4395
Qy 4411 GAGCCAGCAGCAGCTTGCACCTGGAAGGTGGATGAAGGTGCCCAAGGAATTAACAAACGA 4470
Db 4396 GAGCCAGCAGCAGCTTGCACCTGGAAGGTGGATGAAGGTGCCCAAGGAATTAACAAACGA 4455
Qy 4471 GGAAGCAGGCTGGACAGGAAGTACATTTGCTCTGGAGGGATCAAAAGTCTCTCAATTTAT 4530
Db 4456 GGAAGCAGGCTGGACAGGAAGTACATTTGCTCTGGAGGGATCAAAAGTCTCTCAATTTAT 4515
Qy 4531 GACAAATGAAGCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCCCTTCCC 4590
Db 4516 GACAAATGAAGCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCCCTTCCC 4575
Qy 4591 GACGGGATGTATCTATTCATNGTGGCTTGGTGTTCGAACTCGCAAAATACAGCAAA 4650
Db 4576 GACGGGATGTATCTATTCATNGTGGCTTGGTGTTCGAACTCGCAAAATACAGCAAA 4635
Qy 4651 GCA----- 4653
Db 4636 GCAGATGCCCATACATACGTAAGATGGAATCTACCCGACACACCTGCTGGCCCGGG 4695
Qy 4654 ----- 4653
Db 4696 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCCTTA 4755
Qy 4654 ----- 4653
Db 4756 GAATCAGTTTCCAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAATG 4815
Qy 4681 CTTGGAAACTCCCTGTGTAACCTGGAAGGTGATGACCGTCTAGACATGAACGCAAGCTG 4740
Db 4816 CTTGGAAACTCCCTGTGTAACCTGGAAGGTGATGACCGTCTAGACATGAACGCAAGCTG 4875
Qy 4741 CCCTTCACTGACAGGTGTGTTGGTGGGACCGAGAGAGGCTCTACGCCCTGAATGTC 4800
Db 4876 CCCTTCACTGACAGGTGTGTTGGTGGGACCGAGAGAGGCTCTACGCCCTGAATGTC 4935
Qy 4801 TTGAAAACTCCCTAAACCTGCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
Db 4936 TTGAAAACTCCCTAAACCTGCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995
Qy 4861 AAGGACCTGGAGAGCTACTCATATGATAGGAGAGAGAGCGGCACTGTGCTTGTGGAC 4920
Db 4996 AAGGACCTGGAGAGCTACTCATATGATAGGAGAGAGAGCGGCACTGTGCTTGTGGAC 5055
Qy 4921 GTGAGAAAGTGAACAGTCCCTGGCCCGCAGTCCCTGCTGCCAGCCGACATCTCA 4980
Db 5056 GTGAGAAAGTGAACAGTCCCTGGCCCGCAGTCCCTGCTGCCAGCCGACATCTCA 5115
Qy 4981 CCCAACATTTTGAAGCTGCAAGGCTGCCACTTCTTGGGCGAGGCAAGATTGAGAAC 5040
Db 5116 CCCAACATTTTGAAGCTGCAAGGCTGCCACTTCTTGGGCGAGGCAAGATTGAGAAC 5175
Qy 5041 GGGCTGTGATCTGTGACAGCATGCCACCAAGTCTCATTTCTCGCTACACGAAAC 5100
Db 5176 GGGCTGTGATCTGTGACAGCATGCCACCAAGTCTCATTTCTCGCTACACGAAAC 5235
Qy 5101 CTCAGCAAAATACATGTCATCCGAAAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160

Db 5236 CTCAGCAAAATACATGCTCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5295
Qy 5161 TTCAACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTACGAAATTCGACATGAAGCAG 5220
Db 5296 TTCAACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTACGAAATTCGACATGAAGCAG 5355
Qy 5221 TACACGCTCGAGGAAATTCCTGGATGAAGTGAATTCCTTGGCACCCTGTGTTTGGC 5280
Db 5356 TACACGCTCGAGGAAATTCCTGGATGAAGTGAATTCCTTGGCACCCTGTGTTTGGC 5415
Qy 5281 GCCCTTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGAGAG 5340
Db 5416 GCCCTTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGAGAG 5475
Qy 5341 GAGTACTTGTCTGTCTTCCAGGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACCTTAT 5400
Db 5476 GAGTACTTGTCTGTCTTCCAGGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACCTTAT 5535
Qy 5401 CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTGGCCTTTTGCTTACAGAGAACCTTAT 5460
Db 5536 CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTGGCCTTTTGCTTACAGAGAACCTTAT 5595
Qy 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCAGCA 5520
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCAGCA 5655
Qy 5521 GGGACCCCTCGCGAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCCCTTGCATT 5580
Db 5656 GGGACCCCTCGCGAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCCCTTGCATT 5715
Qy 5581 TCCTTCAGGAGCGATTTACTTTGGCGTCTCTCATACAGGATAAATTAAGGTCATTGCTGC 5640
Db 5716 TCCTTCAGGAGCGATTTACTTTGGCGTCTCTCATACAGGATAAATTAAGGTCATTGCTGC 5775
Qy 5641 AAGGAAACCTCTGAGGAGTCCGACATCGGACATCGGACATCGGACATCGGACATCGGAC 5700
Db 5776 AAGGAAACCTCTGAGGAGTCCGACATCGGACATCGGACATCGGACATCGGACATCGGAC 5835
Qy 5701 AG 5702
Db 5836 AG 5837
RESULT 12
US-10-276-774-137
; Sequence 137, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 137
; LENGTH: 3131
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-137
Query Match 39.0%; Score 2564.8; DB 13; Length 3131;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2781; Conservative 0; Mismatches 2; Indels 195; Gaps 2;
Qy 2875 GCACATAGATGAATTCAGCGCAATTTGATGCTCTTCGTACACAGCTCTACTGTAATC 2934


```
QY 4945 GCCAGTCCACCTGCTGCTGCCAGCCCGACATCTCACCAACATTTTGAAGCTGTCAAG 5004
Db 2176 GCCAGTCCACCTGCTGCCAGCCCGACATCTCACCAACATTTTGAAGCTGTCAAG 2235
QY 5005 GGCTGCACCTTTTGGGGAGGCAAGATGAGAACGGGCTCTGCATCTGTGACGCGATG 5064
Db 2236 GGCTGCACCTTTTGGGGAGGCAAGATGAGAACGGGCTCTGCATCTGTGACGCGATG 2295
QY 5065 CCAGCAAGTCTGCTATCTCCGCTCAACAGAAACCTCAGCAAAATACATGCAATCCGAAA 5124
Db 2296 CCAGCAAGTCTGCTATCTCCGCTCAACAGAAACCTCAGCAAAATACATGCAATCCGAAA 2355
QY 5125 GAGATAGACCTCAGAGCCCTGAGCTGTATCCACTTCCAAATACAGTATCTCAAT 5184
Db 2356 GAGATAGACCTCAGAGCCCTGAGCTGTATCCACTTCCAAATACAGTATCTCAAT 2415
QY 5185 GGAACCAATATAATCTACGAATCGACATGAACGATGACGCTCGAGGAATCTCTGGAT 5244
Db 2416 GGAACCAATATAATCTACGAATCGACATGAACGATGACGCTCGAGGAATCTCTGGAT 2475
QY 5245 AAGAATGACCATTCCTTGGCACCCTGCTGTGTTTGGCGCCTCTTCCAAACAGCTTCCCTGTC 5304
Db 2476 AAGAATGACCATTCCTTGGCACCCTGCTGTGTTTGGCGCCTCTTCCAAACAGCTTCCCTGTC 2535
QY 5305 TCAATCGTCAGGTGAACAGCGGAGGCGGAGGAGGATCTTCTGCTGTGTTTCCAGAA 5364
Db 2536 TCAATCGTCAGGTGAACAGCGGAGGCGGAGGAGGATCTTCTGCTGTGTTTCCAGAA 2595
QY 5365 TTGGAGTGTCTGCTGATCTTACGGAACGCTAGCGGACAGAGGATCTCAAGTGGAGT 5424
Db 2596 TTGGAGTGTCTGCTGATCTTACGGAACGCTAGCGGACAGAGGATCTCAAGTGGAGT 2655
QY 5425 CGTTACCTTTGGCTTTGCTTACAGAGAACCTTATCTGTTTGTGACCCACTTCAACTCA 5484
Db 2656 CGTTACCTTTGGCTTTGCTTACAGAGAACCTTATCTGTTTGTGACCCACTTCAACTCA 2715
QY 5485 CTGGAAGTAAATGAGATCAGGACGCTCTCAGCAGGAGGACCCCTGCGGAGCGTACCTG 5544
Db 2716 CTGGAAGTAAATGAGATCAGGACGCTCTCAGCAGGAGGACCCCTGCGGAGCGTACCTG 2775
QY 5545 GAGATCCGAACCCGCGCTACCTGGGCGCTGCAATTCCTCAGGAGCGATTTACTTGGCG 5604
Db 2776 GAGATCCGAACCCGCGCTACCTGGGCGCTGCAATTCCTCAGGAGCGATTTACTTGGCG 2835
QY 5605 TCCTATACAGGATTAATTAAGGTCATTTGCTGCAAGGAAACCTCGTGAAGGATCC 5664
Db 2836 TCCTATACAGGATTAATTAAGGTCATTTGCTGCAAGGAAACCTCGTGAAGGATCC 2895
QY 5665 GGCACCTGAACACCCAGCGGCGCTGCTCCACTCCCGCAG 5702
Db 2896 GGCACCTGAACACCCAGCGGCGCTCCACTCCCGCAG 2933
```

RESULT 13

```
US-10-262-511-7
; Sequence 7, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Mesra
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.

Query Match 36.7%; Score 2414; DB 13; Length 2542;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2475; Conservative 0; Mismatches 5; Indels 48; Gaps 1;

QY 1864 CAAGGGAAGCTTGAAGTGGAGAAATATCGAAACTGGAGAAGATCAATGCTGAGGAGCAG 1923
Db 11 CAGGGAAGCTTGAAGTGGAGAAATATCGAAACTGGAGAAGATCAATGCTGAGGAGCAG 70
QY 1924 CTCAAAATTCAGGAGCTCCCAAGAGAAACTCGAGAAGGCT----- 1962
Db 71 CTCAAAATTCAGGAGCTCCCAAGAGAAACTCGAGAAGGCTTAAAGCCAGCAGGAGGCC 130
QY 1963 -----GCAAGAGAGCGAGCCGAGAGGAGCTTGAAGAG 1995
Db 131 ACCGAGCTGCTGCAGAAATATCCGCCAGGCAAAAGAGCGAGCGAGGAGGAGCTGAGAG 190
QY 1996 CTCGAGAACCGAGAGGATTTCTTCTGAGGAGCTCAGAAAGAGCTGTTGAAGCTGAGGAA 2055
Db 191 CTGCAAGAACCGAGAGGATTTCTTCTGAGGAGCTCAGAAAGAGCTGTTGAAGCTGAGGAA 250
QY 2056 CGCCGCCATTTCTCTGGAGAAACAAGGTAAGAGACTAGAGACCATGAGCGGTAGAGAAAC 2115
Db 251 CGCCGCCATTTCTCTGGAGAAACAAGGTAAGAGACTAGAGACCATGAGCGGTAGAGAAAC 310
```

```
APPLICANT: Anderson, David W.
APPLICANT: Zhong, Mei
APPLICANT: Catterton, Elina
APPLICANT: Ji, Weizhen
APPLICANT: Miller, Charles E.
APPLICANT: Rastelli, Luca
APPLICANT: Stone, David J.
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shimkets, Richard A.
APPLICANT: Rothenberg, Mark E.
APPLICANT: Leach, Martin D.
APPLICANT: Agee, Michele L.
APPLICANT: Berghs, Constance
FILE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-462C
CURRENT APPLICATION NUMBER: US/10/262,511
CURRENT FILING DATE: 2003-05-28
PRIOR APPLICATION NUMBER: 60/326,483
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/373,815
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/381,642
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: 60/381,038
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/373,260
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/373,826
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,435
PRIOR FILING DATE: 2001-10-05
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: CuroSeqlist version 0.1
SEQ ID NO 7
LENGTH: 2542
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(2542)
US-10-262-511-7
```


QY	2116	AGACTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATT	2175
DB	311	AGACTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATT	370
QY	2176	CTGGAGCTCGAAGAGAAACATTCGGGAGGCCAGTCTCAGCCAGCACCTAGAGTGCAC	2235
DB	371	CTGGAGCTCGAAGAGAAACATTCGGGAGGCCAGTCTCAGCCAGCACCTAGAGTGCAC	430
QY	2236	CTGAAACAGAAAGAGCAGCACTATGAGGAAAAATTAAGTGTGTGGACAATCAGATAAAG	2295
DB	431	CTGAAACAGAAAGAGCAGCACTATGAGGAAAAATTAAGTGTGTGGACAATCAGATAAAG	490
QY	2296	AAAGACTTGGCTGACAAAGAGACACTTGGAGAACATGATGCAGAGACACAGGAGGAGGCC	2355
DB	491	AAAGACTTGGCTGACAAAGAGACACTTGGAGAACATGATGCAGAGACACAGGAGGAGGCC	550
QY	2356	CATGAGAAGGGCAAAATTTCTCAGCGAACAGAAAGCGATGATCAATGCTATGGATTCCAAG	2415
DB	551	CATGAGAAGGGCAAAATTTCTCAGCGAACAGAAAGCGATGATCAATGCTATGGATTCCAAG	610
QY	2416	ATCAGATTCCTGTGAACAGAGGATTTGGAACTGTCTGAAGCCAAATAACTTGCAGCAAAAT	2475
DB	611	ATCAGATTCCTGTGAACAGAGGATTTGGAACTGTCTGAAGCCAAATAACTTGCAGCAAAAT	670
QY	2476	AGCAGTCTTTTTTACCACAAAGGAAACATGAAGGCCAACAGAGAGATGATTTCTGAACCTCAGG	2535
DB	671	AGCAGTCTTTTTTACCACAAAGGAAACATGAAGGCCAACAGAGAGATGATTTCTGAACCTCAGG	730
QY	2536	CAACAGAAATTTTACTCTTGGAGACACAGGCTTGGAAATTTGAGGCCAGAACCCGAAACTG	2595
DB	731	CAACAGAAATTTTACTCTTGGAGACACAGGCTTGGAAATTTGAGGCCAGAACCCGAAACTG	790
QY	2596	GAGGAGCGCTGGAGAGATCAGCCACCAAGACACAGTGCACAGAAATCGGCTGCTGGAAC	2655
DB	791	GAGGAGCGCTGGAGAGATCAGCCACCAAGACACAGTGCACAGAAATCGGCTGCTGGAAC	850
QY	2656	CTGGAGACAAAGATTTGGGGAGGTCAGTCTTAGAGCACGAGAGCAGAAACTGGAGCTCAAG	2715
DB	851	CTGGAGACAAAGATTTGGGGAGGTCAGTCTTAGAGCACGAGAGCAGAAACTGGAGCTCAAG	910
QY	2716	CGCCAGCTCACAGAGCTACAGCTCTCCCTCGAGGAGCGCGAGTCAACGTTGACAGCCCTG	2775
DB	911	CGCCAGCTCACAGAGCTACAGCTCTCCCTCGAGGAGCGCGAGTCAACGTTGACAGCCCTG	970
QY	2776	CAGCTCACCGGGCGCCCTTGGAGAGCCAGCTTCGCCAGCGGAAGACAGAGCTTGGAAAGAG	2835
DB	971	CAGCTCACCGGGCGCCCTTGGAGAGCCAGCTTCGCCAGCGGAAGACAGAGCTTGGAAAGAG	1030
QY	2836	ACCACAGCAAGCTGAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAAATCCAG	2895
DB	1031	ACCACAGCAAGCTGAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAAATCCAG	1090
QY	2896	CGCAATTTTGTGATCTTCTTCGTACAGCTGTACTCTTAATCAAGACCTTGGAGGAGCAGCTA	2955
DB	1091	CGCAATTTTGTGATCTTCTTCGTACAGCTGTACTCTTAATCAAGACCTTGGAGGAGCAGCTA	1150
QY	2956	AACAGCTGACCGAGGACAAACGCTGAACTCAACAAACAAAACTTCTACTTGTCCAAACAA	3015
DB	1151	AACAGCTGACCGAGGACAAACGCTGAACTCAACAAACAAAACTTCTACTTGTCCAAACAA	1210
QY	3016	CTCGATGAGGCTTCTTGGCGCCCAACGACGAGATTTGTACAACTGCGAAGTGAAGTGGACCAT	3075
DB	1211	CTCGATGAGGCTTCTTGGCGCCCAACGACGAGATTTGTACAACTGCGAAGTGAAGTGGACCAT	1270
QY	3076	CTCCGCGGGAGATTCAGGAAACGAGAGATGCAGCTTACCAGCCAGAAACGATGGAG	3135
DB	1271	CTCCGCGGGAGATTCAGGAAACGAGAGATGCAGCTTACCAGCCAGAAACGATGGAG	1330
QY	3136	GCTCTGAAGACCAAGTGCAACATGCTGGAGGAAACAGTCTATGATTTGGAGGCCCTTAAC	3195
DB	1331	GCTCTGAAGACCAAGTGCAACATGCTGGAGGAAACAGTCTATGATTTGGAGGCCCTTAAC	1390

Qy	3196	GATGAGCTGCTAGAAAAAGACGGCAGTGGGAGGCTTGAGGAGCGCTTCGGTGGTATGAG	3259
Db	1391	GATGAGCTGCTGAAAAAGACGGCAGTGGGAGGCTTGAGGAGCGCTTCGGTGGTATGAG	1450
Qy	3256	AAATCCCAAGTTTCAGTGTCCGGTTTCAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAG	3315
Db	1451	AAATCCCAAGTTTCAGTGTCCGGTTTCAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAG	1510
Qy	3316	AGCAGGCGGAGAGCGGATCAAGCGGATCACCGAGTCTGCCAGTGTGTGGAGCTGGCAGTG	3375
Db	1511	AGCAGGCGGAGAGCGGATCAAGCGGATCACCGAGTCTGCCAGTGTGTGGAGCTGGCAGTG	1570
Qy	3376	AAGGAGCACAAAGGCTGAGATTCTTCGCTCTCCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAG	3435
Db	1571	AAGGAGCACAAAGGCTGAGATTCTTCGCTCTCCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAG	1630
Qy	3436	GCCGAGAGCTCTCTGTCAAGAGCTCAATGACTCTGGAGAAAGCATGCTGTGTTGAAATG	3495
Db	1631	GCCGAGAGCTCTCTGTCAAGAGCTCAATGACTCTGGAGAAAGCATGCTGTGTTGAAATG	1690
Qy	3496	AATGCCCGAGAGCTTACAGCAGAGAGCTGGAGACTGAACGAGAGCTCAAACAGAGGCTTCTG	3555
Db	1691	AATGCCCGAGAGCTTACAGCAGAGAGCTGGAGACTGAACGAGAGCTCAAACAGAGGCTTCTG	1750
Qy	3556	GAAGAGCAAGCCAAATTTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGGTCTG	3615
Db	1751	GAAGAGCAAGCCAAATTTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGGTCTG	1810
Qy	3616	ACTCAAGAGCTGCAGAAAGCTCTAGATCGGGCTGATCTACTGAAGCAGAGAAAGAGTGAC	3675
Db	1811	ACTCAAGAGCTGCAGAAAGCTCTAGATCGGGCTGATCTACTGAAGCAGAGAAAGAGTGAC	1870
Qy	3676	TTGAGTATCAGCTTGGAAAAATTTCAGGTCTCTATTCTCATGAAAAGGTGAAAATGAA	3735
Db	1871	TTGAGTATCAGCTTGGAAAAATTTCAGGTCTCTATTCTCATGAAAAGGTGAAAATGAA	1930
Qy	3736	GGCACTATTCTCAACAAACCAACTCATTTGATTTCTGCAAGCCAAATTTGGACCAACT	3795
Db	1931	GGCACTATTCTCAACAAACCAACTCATTTGATTTCTGCAAGCCAAATTTGGACCAACT	1990
Qy	3796	GCTAAAAGAAAAAGGTTTATTTAGTCGACGGAAGAGGACCCGCTGTTTACCCACACAG	3855
Db	1991	GCTAAAAGAAAAAGGTTTATTTAGTCGACGGAAGAGGACCCGCTGTTTACCCACACAG	2050
Qy	3856	GTTCTCTGCAGTCAATGAGCTGAAGCTGGCCCTGAGAGGAAGAGAGCTCGTGTGCA	3915
Db	2051	GTTCTCTGCAGTCAATGAGCTGAAGCTGGCCCTGAGAGGAAGAGAGCTCGTGTGCA	2110
Qy	3916	GAGCTAGAGGAAGCCCTTCAGACACCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCT	3975
Db	2111	GAGCTAGAGGAAGCCCTTCAGACACCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCT	2170
Qy	3976	GCCACCGCAAGCAACGGACCCACCCATCCACGCCAGCCACCGGAGCAGCAG	4035
Db	2171	GCCACCGCAAGCAACGGACCCACCCATCCACGCCAGCCACCGGAGCAGCAG	2230
Qy	4036	ATCGGCATGTCCGCATGTCGGTCCGAGAGCACAGCCAGTCCATGAGCTCGTG	4095
Db	2231	ATCGGCATGTGTGCATGTCGGTCCGAGAGCACAGCCAGTCCATGAGCTCGTG	2290
Qy	4096	GCCCGCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTT	4155
Db	2291	GCCCGCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTT	2350
Qy	4156	AAGGAAACGATGCACCAATATTCCTCACCGATTCAAGCTAGSACTGAACATCGGAGCC	4215
Db	2351	AAGGAAACGATGCACCAATATTCCTCACCGATTCAAGCTAGSACTGAACATCGGAGCC	2410
Qy	4216	ACAAGTGTGTGTGTCTGGATACCGTGCACCTTTGGACGCCAGGATCCAAATGTCTC	4275
Db	2411	ACAAGTGTGTGTGTCTGGATACCGTGCACCTTTGGACGCCAGGATCCAAATGTCTC	2470
Qy	4276	GAATGTCAAGTGTATGTGTCAACCCCAAGTGTCTTCCACGCTTTCGACGCCACTCGCGGCTG	4335

||||| 2471 GAATGTAGTGTGTCACCCCAAGTGTCTCCACGTCTCCAGCGCCACTCGCGCTTG 2530
|||||

Qy 4336 CACTCTGA 4343
|||||
Db 2531 CCTGTCGA 2538

RESULT 14

US-10-262-511-5

; Sequence 5, Application US/10262511

; Publication No. US2004003822A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glennnda

; APPLICANT: Millet, Isabelle

; APPLICANT: Peyman, John A.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Ju, Jingfang

; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Patturajan, Meera

; APPLICANT: Svytek, Kimberly A.

; APPLICANT: Edinger, Shomut R.

; APPLICANT: Elierman, Karen

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Ort, Tatiana

; APPLICANT: Gorman, Linda

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Anderson, David W.

; APPLICANT: Zhong, Mei

; APPLICANT: Catterton, Elina

; APPLICANT: Ji, Weizhen

; APPLICANT: Miller, Charles E.

; APPLICANT: Rastelli, Luca

; APPLICANT: Stone, David J.

; APPLICANT: Pena, Carol E. A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Rothenberg, Mark E.

; APPLICANT: Leach, Martin D.

; APPLICANT: Agee, Michele L.

; APPLICANT: Berghs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-462C

; CURRENT APPLICATION NUMBER: US/10/262,511

; PRIOR FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/373,815

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,642

; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/373,260

; PRIOR FILING DATE: 2002-04-17

; PRIOR APPLICATION NUMBER: 60/373,826

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,435

; PRIOR FILING DATE: 2001-10-05

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 439

; SOFTWARE: CuraSeqList version 0.1

; SEQ ID NO 5

; LENGTH: 2497

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(2497)
US-10-262-511-5

Query Match 35.2%; Score 2317.2; DB 13; Length 2497;

Best Local Similarity 96.2%; Pred. No. 0;

Matches 2432; Conservative 0; Mismatches 3; Indels 93; Gaps 2;

Qy 1864 CAAGGGAAGCCTGGAAGTGGGAGAAATATGCGAAATCTGGAGAGATCAATGCTGACAGCAG 1923

Db 11 CAGGGGAAGCCTGGAAGTGGGAGAAATATGCGAAATCTGGAGAGATCAATGCTGACAGCAG 70

Qy 1924 CTCAAAATTCAGGAGCTCCAGAGAAAATCTGGAGAGGCT----- 1962

Db 71 CTCAAAATTCAGGAGCTCCAGAGAAAATCTGGAGAGGCTGTAAAGCCAGCACGAGGCC 130

Qy 1963 -----GCAAAGGAGCGAGCCGAGAGGAGGAGCTGGAGAA 1995

Db 131 ACCGAGCTGCTGCAGAAATATCCGCCAGGCAAAAGGAGCGAGCCGAGAGGAGCTGGAGAA 190

Qy 1996 CTGCAGAAACCGAGAGGATTTCTTGAAGGATATCAGAAAAGAGCTGGTGGAGCTGGAGAA 2055

Db 191 CTGCAGAAACCGAGAGGATTTCTTGAAGGATATCAGAAAAGAGCTGGTGGAGCTGGAGAA 250

Qy 2056 CGCGCCATTTCTTGGAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGTGGAGAAAC 2115

Db 251 CGCGCCATTTCTTGGAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGTGGAGAAAC 310

Qy 2116 AGACTGAAGGATGATCCAGACAAAATCCCAACAGATCCAGCAGATGGTGTGATAAATT 2175

Db 311 AGACTGAAGGATGATCCAGACAAAATCCCAACAGATCCAGCAGATGGTGTGATAAATT 370

Qy 2176 CTGGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCAGACCATAGAGTGCAC 2235

Db 371 CTGGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCAGACCATAGAGTGCAC 430

Qy 2236 CTGAAAACAGAAAGAGCAGCAGTATGAGGAAAAGATTAAAGTGTGGCAATCAGATAAG 2295

Db 431 CTGAAAACAGAAAGAGCAGCAGTATGAGGAAAAGATTAAAGTGTGGCAATCAGATAAG 490

Qy 2296 AAAGACTGCTGACAGGAGACACTGGAGAAACATGATGTCAGAGACACAGAGGAGGCGC 2355

Db 491 AAAGACTGCTGACAGGAGACACTGGAGAAACATGATGTCAGAGACACAGAGGAGGCGC 550

Qy 2356 CATGAGAAAGGCAAAATTTCTCAGGAAACAGAGCCGATGATCAATGCTATGGATTCCAAG 2415

Db 551 CATGAGAAAGGCAAAATTTCTCAGGAAACAGAGCCGATGATCAATGCTATGGATTCCAAG 610

Qy 2416 ATCAGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTGAGCAAAAT 2475

Db 611 ATCAGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTGAGCAAAAT 670

Qy 2476 AGCAGTCTTTTACCCAAAGGAAACATGAGGCCCAAGAGAGATGATTTCTGAACCTCAGG 2535

Db 671 AGCAGTCTTTTACCCAAAGGAAACATGAGGCCCAAGAGAGATGATTTCTGAACCTCAGG 730

Qy 2536 CAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGCCCGAGAAACCGAAAACCTG 2595

Db 731 CAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGCCCGAGAAACCGAAAACCTG 790

Qy 2596 GAGGAGCGCTGGAGAGATCCACCAGAGACACAGTGCACAGATCGGCTGCTGGAA 2655

Db 791 GAGGAGCGCTGGAGAGATCCACCAGAGACACAGTGCACAGATCGGCTGCTGGAA 850

Qy 2656 CTGAGACAAAGATTGGGGAGGTCAGTCTAGAGCACAGAGAGCAGAAAATCTGGAGCTCAAG 2715

Db 851 CTGAGACAAAGATTGGGGAGGTCAGTCTAGAGCACAGAGAGCAGAAAATCTGGAGCTCAAG 910

Qy 2716 CGCCAGCTCACAGACTACAGTCTCTCTGAGGAGCGGAGTTCAGTTGACAGCCCTG 2775

Db 911 CGCCAGCTCACAGACTACAGTCTCTCTGAGGAGCGGAGTTCAGTTGACAGCCCTG 970

2776 CAGGTGACGGGGCCCTGGAGAGCCAGCTTCCGACGGGAGACAGAGCTGGAAGAG 2835
Db
971 CAGGTGACGGGGCCCTGGAGAGCCAGCTTCCGACGGGAGACAGAGCTGGAAGAG 1030
Qy
2836 ACCACAGCAGAAGCTGAAGAGAGATCCAGGACCTCAACGGACATAGAGATGAATCCAG 2895
Db
1031 ACCACAGCAGAAGCTGAAGAGAGATCCAGGACCTCAACGGACATAGAGATGAATCCAG 1090
Qy
2896 GCGAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCAACAGACCTCGAGGAGCAGCTA 2955
Db
1091 GCGAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCAACAGACCTCGAGGAGCAGCTA 1150
Qy
2956 AACCAAGCTGACCGAGGACAAACGCTGAACCTCAACAAACCAAACTTCTACTTTGCCAAACA 3015
Db
1151 AACCAAGCTGACCGAGGACAAACGCTGAACCTCAACAAACCAAACTTCTACTTTGCCAAACA 1210
Qy
3016 CTCGATGAGCTTCTGGGCGCCAAACGAGATGTACAACCTGCGAAGTGAAGTGAACCAT 3075
Db
1211 CTCGATGAGCTTCTGGGCGCCAAACGAGATGTACAACCTGCGAAGTGAAGTGAACCAT 1270
Qy
3076 CTCGCGCGGAGATCACGGAAACGAGATGCAAGCTTACCAGCCAGAAACCAACGATGGAG 3135
Db
1271 CTCGCGCGGAGATCACGGAAACGAGATGCAAGCTTACCAGCCAGAAACCAACGATGGAG 1330
Qy
3136 GCTCTGAAGACCAAGTGCACCATGCTGGAGGAAACAGGTCAAGATTTGGAGCCCTTAAC 3195
Db
1331 GCTCTGAAGACCAAGTGCACCATGCTGGAGGAAACAGGTCAAGATTTGGAGCCCTTAAC 1390
Qy
3196 GATGAGCTCTAGAAAAGAGCGGAGTGGAGGCTGGAGAGCGTCTCGGGTCAATGAG 3255
Db
1391 GATGAGCTCTAGAAAAGAGCGGAGTGGAGGCTGGAGAGCGTCTCGGGTCAATGAG 1450
Qy
3256 AAATCCCAAGTTTGAAGTGTCCGGTTCCAGAGCTGCAAGAAATCTCGAGCAACGAGAAAACAG 3315
Db
1451 AAATCCCAAGTTTGAAGTGTCCGGTTCCAGAGCTGCAAGAAATCTCGAGCAACGAGAAAACAG 1510
Qy
3316 AGCAGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCAGGCTGGAGCTGCAGTG 3375
Db
1511 AGCAGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCAGGCTGGAGCTGCAGTG 1570
Qy
3376 AAGGAGCAAGAGCTGAGATTTCTCGTCTGCAGCAGGCTCTCAAAAGAGCAGAAGCTGAAG 3435
Db
1571 AAGGAGCAAGAGCTGAGATTTCTCGTCTGCAGCAGGCTCTCAAAAGAGCAGAAGCTGAAG 1630
Qy
3436 GCGAGAGCTCTCTGCAAGCTCAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3495
Db
1631 GCGAGAGCTCTCTGCAAGCTCAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1690
Qy
3496 AATGCCCGAAGCTTACAGCAGAGCTGGAGACTGAACAGAGAGCTCAAAAGAGGCTTCTG 3555
Db
1691 AATGCCCGAAGCTTACAGCAGAGCTGGAGACTGAACAGAGAGCTCAAAAGAGGCTTCTG 1750
Qy
3556 GAAGAGCAAGCCAAATTAACAGCAGCAGATGGACCTGCAGAAAATCAATTTCCGCTG 3615
Db
1751 GAAGAGCAAGCCAAATTAACAGCAGCAGATGGACCTGCAGAAAATCAATTTCCGCTG 1810
Qy
3616 ACTCAAGGACTGAAGAGCTTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGAAC 3675
Db
1811 ACTCAAGGACTGAAGAGCTTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGAAC 1870
Qy
3676 TTGGAGTATCAGCTGGAAGAAATTCATGAGTTCTTCTATTCATGAAAAGGTGAAAATGGAA 3735
Db
1871 TTGGAGTATCAGCTGGAAGAAATTCATGAGTTCTTCTATTCATGAAAAGGTGAAAATGGAA 1930
Qy
3736 GGCATATTTCTCAACAAACCAATTCATGATTTCTGCAAGCCAAATGACCAACT 3795
Db
1931 GGCATATTTCTCAACAAACCAATTCATGATTTCTGCAAGCCAAATGACCAACT 1990
Qy
3796 GCTAAAAGAAAAGGGTTTATTTAGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3855
Db
1991 GCTAAAAGAAAAGGGTTTATTTAGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2005
Qy
3856 GTTCTCTGAGTACATGAGCTGAAGCTGGCCCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3915

2006 GTTCTCTGAGTACATGAGCTGAAGCTGGCCCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2065
Qy
3916 GAGCTAGAGAGCCCTTTCAGAGAGCCGATCGAGCTCCGGTCCGCCCGGAGGAGGAGGAGGAGGAG 3975
Db
2066 GAGCTAGAGAGCCCTTTCAGAGAGCCGATCGAGCTCCGGTCCGCCCGGAGGAGGAGGAGGAGGAG 2125
Qy
3976 GCCACCGCAAGCAACGAGCAGCCACCAATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4035
Db
2126 GCCACCGCAAGCAACGAGCAGCCACCAATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2185
Qy
4036 ATGCCATGTCGCCATCGTGGCTCGCAGAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4095
Db
2186 ATGCCATGTCGCCATCGTGGCTCGCAGAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2245
Qy
4096 GCCCGCCATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4155
Db
2246 GCCCGCCATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2305
Qy
4156 AAGAACGATGACACCAATATTCACCGATTCACCGATTCACCGATTCACCGATTCACCGATTCACCGATTC 4215
Db
2306 AAGAACGATGACACCAATATTCACCGATTCACCGATTCACCGATTCACCGATTCACCGATTCACCGATTC 2365
Qy
4216 ACAAAGT 4275
Db
2366 ACAAAGT 2425
Qy
4276 GAATGTGAGT 4335
Db
2426 GAATGTGAGT 2485
Qy
4336 CCTGCTGA 4343
Db
2486 CCTGCTGA 2493

RESULT 15

US-10-262-511-3
; Sequence 3, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerkhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Carterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511

```
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curasequid version 0.1
; SEQ ID NO 3
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1870)
; US-10-262-511-3

Query Match
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1849; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 6 CGCAGTGGGGAGATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTTGATGCTGTGC 65
DB 1 CACCGGTACCAACCATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTTGATGCTGTGC 60
QY 66 TGCTGAACCATGTCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACCCACC 125
DB 61 TGCTGAACCATGTCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACCCACC 120
QY 126 CTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTT 185
DB 121 CTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTT 180
QY 186 TGTCTCTTTGAGAAATGCAATGCTGCTCTGATGAAGATTAAGCAGTGAACAATTT 245
DB 181 TGTCTCTTTGAGAAATGCAATGCTGCTCTGATGAAGATTAAGCAGTGAACAATTT 240
QY 246 TGTCCGGAAGTATTCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGA 305
DB 241 TGTCCGGAAGTATTCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGA 300
QY 306 CTTCCGAGTCAAGATCTTTGAGTTGCTGCTCTGATGAAGTGAAGTGAAGTGAAG 365
DB 301 CTTCCGAGTCAAGATCTTTGAGTTGCTGCTCTGATGAAGTGAAGTGAAGTGAAG 360
QY 366 GAAAGCAACCGGGACATCTATGCTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 425
DB 361 GAAAGCAACCGGGACATCTATGCTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
QY 426 GGAGCAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAAGCAAGCCCTGT 485
DB 421 GGAGCAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAAGCAAGCCCTGT 480
QY 486 GATCCCCCAATACAGTATGCTTTTCAGACAAAAATCACCTTTATCTGATGGAGGATA 545
DB 481 GATCCCCCAATACAGTATGCTTTTCAGACAAAAATCACCTTTATCTGATGGAGGATA 540

546 TCAGCTCGAGGGGACTTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAA 605
541 TCAGCTCGAGGGGACTTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAA 600
606 CTTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCAACGCGTTCACTGATGGG 665
601 CTTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCAACGCGTTCACTGATGGG 660
666 ATACGTGATCGAGACATCAAGCCTTGAGAACATTTCTGTTGACCGCACAGACACATCAA 725
661 ATACGTGATCGAGACATCAAGCCTTGAGAACATTTCTGTTGACCGCACAGACACATCAA 720
726 GCTGCTGATTTTGGATCTCGCGGGAATGAATTTCAACAGAGATGTTGAATGCCAACT 785
721 GCTGCTGATTTTGGATCTCGCGGGAATGAATTTCAACAGAGATGTTGAATGCCAACT 780
786 CCCGATTGGGACCCCGAGATTACATGGCTTCTGAAAGTGTGACTGTGATGAACGGGGATGG 845
781 CCCGATTGGGACCCCGAGATTACATGGCTTCTGAAAGTGTGACTGTGATGAACGGGGATGG 840
846 AAAAGGCACTACGGCCTGGACTGTGCTGTGTGCTGAGTGGGCGTGTATGCTATAGAT 905
841 AAAAGGCACTACGGCCTGGACTGTGCTGTGTGCTGAGTGGGCGTGTATGCTATAGAT 900
906 GATTTATGGGAGATCCCCCTTCGCGAGAGGAACTCTGCCAGAACCTTTCATACATTAAT 965
901 GATTTATGGGAGATCCCCCTTCGCGAGAGGAACTCTGCCAGAACCTTTCATACATTAAT 960
966 GATTTATGGGAGATCCCCCTTCGCGAGAGGAACTCTGCCAGAACCTTTCATACATTAAT 1025
961 GATTTATGGGAGATCCCCCTTCGCGAGAGGAACTCTGCCAGAACCTTTCATACATTAAT 1020
1026 TCTGATTCAGAGCTTGTGTCGCGCCAGAGAGAGACTGAACTTTTGAAGTCTTTGCTG 1085
1021 TCTGATTCAGAGCTTGTGTCGCGCCAGAGAGAGACTGAACTTTTGAAGTCTTTGCTG 1080
1086 CCATCTCTTCTCTAAAATTTGACTGGAACCAATTCGTAACCTCTCTCCCCCTTCGT 1145
1081 CCATCTCTTCTCTAAAATTTGACTGGAACCAATTCGTAACCTCTCTCCCCCTTCGT 1140
1146 TCCACACCTCAAGTCTGACATGACACTCCAATTTTGAATGAACAGAGAGATTCGTG 1205
1141 TCCACACCTCAAGTCTGACATGACACTCCAATTTTGAATGAACAGAGAGATTCGTG 1200
1206 GGTTCATCTCTCCGTCGAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGT 1265
1201 GGTTCATCTCTCCGTCGAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGT 1260
1266 TGTGGGTTTTGTAAGAGAGCACTGGGATTTCTTGGTATCTGAGTCTGTTGTGTC 1325
1261 TGTGGGTTTTGTAAGAGAGCACTGGGATTTCTTGGTATCTGAGTCTGTTGTGTC 1320
1326 GGTTCGACTCCCTTCGCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAA 1385
1321 GGTTCGACTCCCTTCGCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAA 1380
1386 AGAGCTACAAGATCTCAGGACAAAGTGTCAAAAGATGGAGAGAAATGACCCGGTTACA 1445
1381 AGAGCTACAAGATCTCAGGACAAAGTGTCAAAAGATGGAGAGAAATGACCCGGTTACA 1440
1446 TCGGAGAGTCTCAGAGTGGAGCTGTGCTTAGTCAAGAGGAGGTGAGCTGAGGCTTC 1505
1441 TCGGAGAGTCTCAGAGTGGAGCTGTGCTTAGTCAAGAGGAGGTGAGCTGAGGCTTC 1500
1506 TGAGACTCAGAGATCCCTCTGAGGAGGACCTTGTCTACCTACATCAGAGATGCAAGTAG 1565
1501 TGAGACTCAGAGATCCCTCTGAGGAGGACCTTGTCTACCTACATCAGAGATGCAAGTAG 1560
1566 CTTAAAGCGAAGTTTGGAGCAAGCTGAGGTTGCTCCAGGAGGATGCAAGAGCACT 1625
1561 CTTAAAGCGAAGTTTGGAGCAAGCTGAGGTTGCTCCAGGAGGATGCAAGAGCACT 1620
1626 GCAGCTTCTCTATGATATCAGAGAGCAGAGCGGAGGCTTCCAGAAATCAAAGAGCAGGA 1685
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:08:00 ; Search time 100 Seconds

(without alignments)

6477.589 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10490

Sequence: 1 MLKFKYGNFLDAGAEPI.....QLNGEIRQVEKSVLRDTC 2053

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9890	94.3	2027	4	Q86UQ9	Q86UQ9 homo sapien
2	9656	92.0	2055	11	Q88938	Q88938 mus musculus
3	7717.5	73.6	1641	11	Q88528	Q88528 mus musculus
4	7521	71.7	1618	11	Q9QX19	Q9QX19 rattus norv
5	4819	45.9	1032	11	Q8C1J1	Q8C1J1 mus musculus
6	2165.5	20.6	494	11	Q88937	Q88937 mus musculus
7	2150.5	20.5	448	11	Q88527	Q88527 rattus norv
8	1839.5	17.5	1854	5	Q9VTY8	Q9VTY8 drosophila
9	1450	13.8	1732	11	Q54874	Q54874 rattus norv
10	1421.5	13.6	1719	4	Q86XX3	Q86XX3 homo sapien
11	1385.5	13.2	1760	4	Q9ULU5	Q9ULU5 homo sapien
12	1382.5	13.2	1711	4	Q9Y5S2	Q9Y5S2 homo sapien
13	1373	13.1	1713	11	Q7TT49	Q7TT49 rattus norv
14	1372	13.1	1702	11	O54875	O54875 rattus norv
15	1363	13.0	1713	11	Q7TT50	Q7TT50 mus musculus
16	1353	12.9	1638	4	Q86XX2	Q86XX2 homo sapien

ALIGNMENTS

RESULT 1

Q86UQ9 ID Q86UQ9 PRELIMINARY; PRT; 2027 AA.
AC Q86UQ9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rho/rac-interacting citron kinase.
GN CIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.Q., Wu S.L., Shan Y.X., Liu S., Xiao P.J.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY257469; AAP13528.1; -
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0005083; P:small GTPase regulatory/interacting protein. . ; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR002219; DAG_PE-Bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000861; REM_repeat.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG_PE-Bind; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.

Q81WQ7 homo sapien
O01583 caenorhabdi
P70335 mus musculus
Q63644 rattus norv
Q28021 bos taurus
Q13464 homo sapien
Q90Y37 brachydanio
Q9UQN5 homo sapien
O75116 homo sapien
P70336 mus musculus
O77819 oryctolagus
O73732 xenopus lae
Q62868 rattus norv
O44368 drosophila
Q9W1B0 drosophila
Q9U779 drosophila
Q9VXE3 drosophila
Q961D4 drosophila
P92199 caenorhabdi
Q988N6 gallus gall
Q8AVM0 xenopus lae
Q86XZ8 homo sapien
Q99646 homo sapien
O00565 homo sapien
Q8SS10 encephalito
Q988N5 gallus gall
Q98A79 arabidopsis
Q91W99 mesembryant

```

DR Pfam: PF00433; pkinase C; 1.
DR ProDom: PD000001; ProC_kinase; 1.
DR SMART: SMC0109; Cl; 1.
DR SMART: SMC0036; CNH; 1.
DR SMART: SMC0233; PH; 1.
DR SMART: SMC0320; S TKC; 1.
DR SMART: SMC0133; S TKC; 1.
DR SMART: SMC0219; TY-KG; 1.
DR PROSITE: PS00867; CPSASE_2; 1.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE: PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE: PS00003; PH DOMAIN; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR KX
SQ SEQUENCE 2027 AA; 231429 MW; 6B1D8C3F661F357B CRC64;

Query Match 94.3%; Score 9890; DB 4; Length 2027;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1954; Conservative 2; Mismatches 4; Indels 106; Gaps 3;

QY 1 MLKFKYGARNPLDAGAAEPIASASRLNLFPOGKPPFMTQQQMSPLSREGILDALFVLFE 60
DB 1 MLKFKYGARNPLDAGAAEPIASASRLNLFPOGKPPFMTQQQMSPLSREGILDALFVLFE 60
QY 61 ECQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATG 120
DB 61 ECQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVVRKATG 120
QY 121 DIYAMKMKKALLAQEQVSFFPEERNILSRSTSPWIPOLQYAFQDQXNHYLMEEYQPGG 180
DB 121 DIYAMKMKKALLAQEQVSFFPEERNILSRSTSPWIPOLQYAFQDQXNHYLMEEYQPGG 180
QY 181 DLASLNRYEDQDENLIOPYLAELTAVSHVLMGYVHRDIKPENILVDRGHKLVDF 240
DB 181 DLASLNRYEDQDENLIOPYLAELTAVSHVLMGYVHRDIKPENILVDRGHKLVDF 240
QY 241 GSAAKNNSNMVNAKLPIGTPTYMAPEVLTVMNGDGKGTGYLDCDWWSVGVIAYEM1YGR 300
DB 241 GSAAKNNSNMVNAKLPIGTPTYMAPEVLTVMNGDGKGTGYLDCDWWSVGVIAYEM1YGR 300
QY 301 SPAEGTSARTFNINWFORFLKFPDDPKVSSDPLDIQSLCGQKERLKFGCLCHPFF 360
DB 301 SPAEGTSARTFNINWFORFLKFPDDPKVSSDPLDIQSLCGQKERLKFGCLCHPFF 360
QY 361 SKIDMNNIRNSPPFPVPTLKSDDDTGNFDEPEKNSVSSPCQLSPGSGEELPFVGFS 420
DB 361 SKIDMNNIRNSPPFPVPTLKSDDDTGNFDEPEKNSVSSPCQLSPGSGEELPFVGFS 420
QY 421 YKALGILGRSESVVGLSPAKTSSMEKKLIIKSKELQDSQDKCHKMEQEMTRLHRVS 480
DB 421 YKALGILGRSESVVGLSPAKTSSMEKKLIIKSKELQDSQDKCHKMEQEMTRLHRVS 480
QY 481 EYEAVALSQEVELKASFTQSLLEQDLATYITECSSLKRSQARMEVSOEDDKALQLLH 540
DB 481 EYEAVALSQEVELKASFTQSLLEQDLATYITECSSLKRSQARMEVSOEDDKALQLLH 540
QY 541 DIPEQRKLOEIKQEQYQAOQVEMRLMMNQLBEDIVSARRRDLYESELRESRLAAEFK 600
DB 541 DIPEQRKLOEIKQEQYQAOQVEMRLMMNQLBEDIVSARRRDLYESELRESRLAAEFK 600
QY 601 RKATECOHLLKAKDQKQEVGEYAKLEKINAEQQLKIQELQSKLEKAVKASTEATELLQ 660
DB 601 RKATECOHLLKAKDQKQEVGEYAKLEKINAEQQLKIQELQSKLEKAVKASTEATELLQ 660
QY 649 ----AKERAARELEKIQNREDSSEGIRKKLVFAEBERRHSLNKNVRLMTERENRLKDD 704
DB 661 NIFQAKERAARELEKIQNREDSSEGIRKKLVFAE----- 694
QY 705 IQTKSQIQQADKILVLEKREHQAQVSAHLVHLKQEQHYEBKIKVLDNQIKKDLAD 764
DB 695 -----ELEKREHQAQVSAHLVHLKQEQHYEBKIKVLDNQIKKDLAD 738

```


QY 1795 RTDCLKWSRLPLAFAYREPVLFWTHNSLEVIQARSSAGTAPARAYLDIPNPRYLGPAL 1854
Db 1819 RTDCLKWSRLPLAFAYREPVLFWTHNSLEVIQARSSAGTAPARAYLDIPNPRYLGPAL 1878
QY 1855 SSGAIYASSYQDKLVI CKGNLVKESGTEHHRGPTSRSSPKNGPPTYNHITKRYA 1914
Db 1879 SSGAIYASSYQDKLVI CKGNLVKESGTEHHRGPTSRSSPKNGPPTYNHITKRYA 1938
QY 1915 SSPAPGEGSHPREPSTPHRYREGRTLRDXSPGRPLERKSPGRMLSTRERSPGRLF 1974
Db 1939 SSPAPGEGSHPREPSTPHRYREGRTLRDXSPGRPLERKSPGRMLSTRERSPGRLF 1998
QY 1975 EDSSGRLLPAGAVRTPLSQVKNKGQSA 2002
Db 1999 EDSSGRLLPAGAVRTPLSQVKNKWDQSS 2026

RESULT 2
C88938
ID C88938 PRELIMINARY; PRT; 2055 AA.
AC C88938;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Rho/rac-interacting citron kinase.
GN CIT OR CRK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=99009084; PubMed=9792683;
RA Di Cunto F., Calautti E., Hsiao J., Ong L., Topley G., Turco E.,
RT Dotto G.P.;
RT "Citron rho-interacting kinase, a novel tissue-specific Ser/Thr kinase
encompassing the rho-rac-binding protein citron.";
J. Biol. Chem. 273:29706-29711(1998).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -I- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; AF086824; AAC72823.1; -;
DR HSP; P00517; ISTD.
DR MGD; MGI:105313; Cit.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR005479; Case L D2.
DR InterPro; IPR002219; DAG_PE-Bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000861; REM repeat.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase C; 1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00133; S_TK_X; 1.

DR PROSITE; PS00867; CPSASE 2; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM 1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_2; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 2055 AA; 235481 MW; 2120CB5E454DA940 CRC64;

Query Match 92.0%; Score 9656; DB 11; Length 2055;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;

QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFTQOQMSPLSREGILDALFALFE 60
Db 1 MLKFKYGVNPPEASASEPIASRASRLNLFQGGKPPFTQOQMSPLSREGILDALFALFE 60

QY 61 ECSOPALMKIKHVNFRVFKYSDTTIAELQIOPSADFEVRSLVCCGHFAEVQVVRKATG 120
Db 61 ECSOPALMKIKHVNFRVFKYSDTTIAELQIOPSADFEVRSLVCCGHFAEVQVVRKATG 120

QY 121 DIYAMKVMKKKALLAQOVQVFFEEERNILSRSTSPWIPQIQYAFQDKNHLVMEYQPGG 180
Db 121 DIYAMKVMKKKALLAQOVQVFFEEERNILSRSTSPWIPQIQYAFQDKNHLVMEYQPGG 180

QY 181 DLLSLNRYSDQDENLQFYLAELILAVSHVMGYVHRDIKPNILVDRTGHIKLVDF 240
Db 181 DFLSLNRYSDQDENLQFYLAELILAVSHVMGYVHRDIKPNILVDRTGHIKLVDF 240

QY 241 GSAAKMNSKMNNAKLPIGTVDYNAPEVLTVMGDKGKTYGLDCDMWSVGVIAYEMLYGR 300
Db 241 GSAAKMNSK--VDAKLPIGTVDYNAPEVLTVMNEDREGTYGLDCDMWSVGVIAYEMLYGK 299

QY 301 SPFAEGTSARTNNINMNFQFLKFPDDPKVSSDFDLIOQLLCOQERLKFEGICCHPFF 360
Db 300 TPTTEGTSARTNNINMNFQFLKFPDDPKVSSDFDLIOQLLCOQERLKFEGICCHPFF 359

QY 361 SKDWMNIRNSPPFPVTLKSDDDTSNFDPEPKNSWSSPCQLSPGSGEHELPPVGFPS 420
Db 360 ARTDWMNIRNSPPFPVTLKSDDDTSNFDPEPKNSWSSPCQLSPGSGEHELPPVGFPS 419

QY 421 YSKALGILGSESVWSGLDSPAKTSMEKKLLIKSKELQDSQDKHCKHWEEMTLHRVS 480
Db 420 YSKALGILGSESVWSGLDSPAKTSMEKKLLIKSKELQDSQDKHCKHWEEMTLHRVS 479

QY 481 EYEAVALSQKVELKASGTQSLLEQDLATYITCSSLKRSLEQARMEVSDQDKALQLLH 540
Db 480 EYEAVALSQKVELKASGTQSLLEQDLATYITCSSLKRSLEQARMEVSDQDKALQLLH 539

QY 541 DIREQSKLOEIKEQEQVQAOVEEMLMNMLEDLVSARRSDLYSELESESLAAEEFK 600
Db 540 DIREQSKLOEIKEQEQVQAOVEEMLMNMLEDLVSARRSDLYSELESESLAAEEFK 599

QY 601 RKATECOHKLKARDQKPEVGEYAKLEKINAEQQLKIQLQEKLEKA----- 648
Db 600 RKANECQHKLMAKADQKPEVGEYSKLEKINAEQQLKIQLQEKLEKAVKASTEATELLQ 659

QY 649 ----AKERAERELEKLNREDSDSEGTRKLVAEERERHLENKVKLETMERENLKDD 704
Db 660 NIRQAXERAERELEKLNREDSDSEGTRKLVAEERERHLENKVKLETMERENLKDD 719

QY 705 IQTKSQIQOQADKILEBEKHREAOVSQAHLVHLKQEQHYEEKIKVLDNQIKKDLAD 764
Db 720 IQTKSQIQOQADKILEBEKHREAOVSQAHLVHLKQEQHYEEKIKVLDNQIKKDLAD 779

QY 765 KETLENMQHBEAHEKGIKLEQKAMINAMDSKIRSLERQIVELSEANKLAANSLFT 824
Db 780 KESLENMQHBEAHEKGIKLEQKAMINAMDSKIRSLERQIVELSEANKLAANSLFT 839

QY 825 QRNNKAQEEVISLROOKFYLETCAGKLAQNRKLEQLEKISHQDHSKRNLELETRL 884
Db 840 QRNNKAQEEVISLROOKFYLETCAGKLAQNRKLEQLEKISHQDHSKRNLELETRL 899

885 REVSLEHESQKLEKQKLTQLOI:SLQERESQITAIQARAALLESQI:ROAKTELETTAAE 944
890 REVSLEHESQKLEKQKLTQLOI:SLQERESQITAIQARAALLESQI:ROAKTELETTAAE 959
945 EBEIQALTARDEIOKQFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKQLEDEAS 1004
960 EBEIQALTARDEIOKQFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKQLEDEAS 1019
1005 GANDEIVQLRSEVDHLRRIITRETVOLTSQKQTMALKTTCNTMLEEQVMDLEALNDELLE 1064
1020 GANDEIVQLRSEVDHLRRIITRETVOLTSQKQTMALKTTCNTMLEEQVMDLEALNDELLE 1079
1065 KERQEWASVVLGDEKQSEFCRVRELQRLMDTEKOSRARDQRIITESHQVVELAVKEHKA 1124
1080 KERQEWASVVLGDEKQSEFCRVRELQRLMDTEKOSRARDQRIITESHQVVELAVKEHKA 1139
1125 EILALQALKEQKLAESLSDKLNLEKKGHAMLENNARSLOQKLETERELKORLLEEQAK 1184
1140 EILALQALKEQKLAESLSDKLNLEKKGHAMLENNARSLOQKLETERELKORLLEEQAK 1199
1185 LOQOMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENTQVLYSHKVKMGTTISQ 1244
1200 LOQOMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENTQVLYSHKVKMGTTISQ 1259
1245 QTKLIDFLQAKMDQPAKKGKLFRRKEDPALPTQVPLQYNELKLAKLEKARCAELSEA 1304
1260 QTKLIDFLQAKMDQPAKKGKLFRRKEDPALPTQVPLQYNELKLAKLEKARCAELSEA 1304
1305 LQKTRIELRSARBEAAHRAKATDHPHSTPATARQIAMSIVRSPEHPSANSLAPSS 1364
1305 LQKTRIELRSARBEAAHRAKATDHPHSTPATARQIAMSIVRSPEHPSANSLAPSS 1364
1365 RKESSTPEFRRRLKERMHNT:PHFNVGLNMRATKCAVCLDTHVFGQASKCLEQVM 1424
1365 RKESSTPEFRRRLKERMHNT:PHFNVGLNMRATKCAVCLDTHVFGQASKCLEQVM 1424
1425 CHPKSTCLPATCGLPAEYATHEAFCDKMNPSGLQKPESSSLHLEGWKVPNNKR 1484
1425 CHPKSTCLPATCGLPAEYATHEAFCDKMNPSGLQKPESSSLHLEGWKVPNNKR 1484
1485 GQGWDRKYIVLGSQKVLIIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544
1485 GQGWDRKYIVLGSQKVLIIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544
1545 A-----EKAEADAKL 1554
1545 ADVPYILKMBESHPTTTCWQRTLVLLAPSPDKQKRWVTALESVVAGGRVSRKAEADAKL 1604
1555 LGNSLLKLEGGDRLDNMCTLPFSDQVVLVGTESGLYALNVLKNSLTHVPCIGAVFOIYII 1614
1605 LGNSLLKLEGGDRLDNMCTLPFSDQVVLVGTESGLYALNVLKNSLTHVPCIGAVFOIYII 1664
1615 KLEKLMIAGERALCLVDVKKVQKSLAQSHLPAQPDISPNI:FEAVKGCFLFGAGKIEK 1674
1665 KLEKLMIAGERALCLVDVKKVQKSLAQSHLPAQPDISPNI:FEAVKGCFLFGAGKIEK 1724
1675 GLCICAMPKQKVLIRNENLSKYIKRKEITETSEPCSHFTNYSILIGTNKRYEIDMKQ 1734
1725 SLCTICAMPKQKVLIRNENLSKYIKRKEITETSEPCSHFTNYSILIGTNKRYEIDMKQ 1784
1735 YTLDEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHEFGVFDVSYGRS 1794
1785 YTLDEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHEFGVFDVSYGRS 1844
1795 RTDDLKWSRLPLAFAYREPVLVFTVFNLSLEVIEIQAPSSAGTAPAYLDIPNRYLGPAL 1854
1845 RTDDLKWSRLPLAFAYREPVLVFTVFNLSLEVIEIQAPSSAGTAPAYLDIPNRYLGPAL 1904
1855 SSGAIYVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSNKGPPTYNEHITKRV 1914
1905 SSGAIYVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSNKGPPTYNEHITKRV 1964

QY 1915 SSAPPEGSPHREPSTPHRY--REGRTELRRDKSPGRPLERKSPGRMLSTRERSPGR 1972
Db 1965 SSAPPEGSPHREPSTPHRYDRGRTELRRDKSPGRPLERKSPGRMLSTRERSPGR 2024
QY 1973 LFDSSSRCLPAGAVRTP:LSOVNKGQSGA 2002
Db 2025 LFDSSSRCLPAGAVRTP:LSOVNKGQSGA 2054
RESULT 3
Q88528 PRELIMINARY; PRT; 1641 AA.
AC O88528;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Citron-K kinase (Fragment).
GN CIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=96128238; PubMed=8543060;
RA Madaule P., Furiyashiki T., Reid T., Ishizaki T., Watanabe G.,
RA Morii N., Narumiya S.;
RT "A novel partner for the GTP-bound forms of rho and rac.";
RT FEBS Lett. 377:243-248(1995).
EN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Madaule P., Eda M., Watanabe N., Fujisawa K., Matsuoka T., Bito H.,
RA Ishizaki T., Narumiya S.;
RT "Role of Citron kinase as a target of the small GTPase Rho in
RT cytokinesis.";
RL Nature 0:0-0(1998).
CC -1. SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; AF070066; AAC27933.1; .
DR MG; MG1:105313; Cit.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein.; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR005479; Cbase_L_D2.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00133; S-TX_X; 1.
DR PROSITE; PS00867; CFSASE_2; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW Kinase.
FT NON_TER
SQ SEQUENCE 1 187659 MN; 6723E20ECA3F22AB CRC64;
Query Match 73.6%; Score 7717.5; DB 11; Length 1641;
Best Local Similarity 90.8%; Pred. No. 3.2e-316;
Matches 1541; Conservative 15; Mismatches 16; Indels 125; Gaps 5;

QY	374	PFVPTLKSDDDTSNFDPEPKNSWVSSPCQLSPGSGBELPFVGFYSKALGILGRSES	433
Db	1	PFVPTLKSDDDTSNFDPEPKNSWVSSVQCLSPGSGBELPFVGFYSKALGILGRSES	60
QY	434	VVSGLDSPAKTSMEKLLIKSKELQSDQKCHQWEOENTRLHRRVSEAVLSQKEVEL	493
Db	61	VVSLDSPAIVSMKLLIKSKELQSDQKCHQWEOENTRLHRRVSEAVLSQKEVEL	120
QY	494	KASQTSRLLEDLATYITTCSSLSKSLRSLQARMESQEDDKALQLLHDIRESQRLQETIK	553
Db	121	KASQTSRLLEDLATYITTCSSLSKSLRSLQARMESQEDDKALQLLHDIRESQRLQETIK	180
QY	554	EOYQAOQVEEMLMNOLBEDIVSARRSDIVSELSRESRLAABEFKPKATCQHKLLKA	613
Db	181	EOYQAOQVEEMLMNOLBEDIVSARRSDIVSELSRESRLAABEFKPKATCQHKLLKA	240
QY	614	KQQGEVGEYAKLEKINAEQQLKIQELOEKLEKA-----AKERAEREL	657
Db	241	KQQGEVGEYSKLEKINAEQQLKIQELOEKLEKAVKASTEATELLQNIROKERAEREL	300
QY	658	EKLQNRDSSGIRKVLVAERRRHSLEKNVRLTMRERENRLKDDIQTKSQIQQVAD	717
Db	301	EKLHNRDSSGIRKVLVAERRRHSLEKNVRLTMRERENRLKDDIQTKSQIQQVAD	321
QY	718	KILELEKHREQAQVSAQHLVHLKQEQHYEBKIKVLDNQIKKDLADKETLENMQRHEE	777
Db	322	---ELEEKHREQAQVSAQHLVHLKQEQHYEBKIKVLDNQIKKDLADKETLENMQRHEE	378
QY	778	EAHEKGKILSEQAMINAMDSKIRSLQRIVELSEANKLAANSSLFTORNKAQEMISE	837
Db	379	EAHEKGKILSEQAMINAMDSKIRSLQRIVELSEANKLAANSSLFTORNKAQEMISE	438
QY	838	LQOQFYLETOAGKLEAQNQKLEEOLEKISHODHDKNRLLEFRLREVSLSHEEQKLE	897
Db	439	LQOQFYLETOAGKLEAQNQKLEEOLEKISHODHDKNRLLEFRLREVSLSHEEQKLE	498
QY	898	LKRQTELEQLSLQERESQUTALQAARAALESQRLQAKTELETTAAEETEQALTARDE	957
Db	499	LKRQTELEQLSLQERESQUTALQAARAALESQRLQAKTELETTAAEETEQALTARDE	558
QY	958	IQRKFDALNSCTVITDLEQNLQNTEDNAELNNQNFYLSKOLDASGANDEIVOLRSEV	1017
Db	559	IQRKFDALNSCTVITDLEQNLQNTEDNAELNNQNFYLSKOLDASGANDEIVOLRSEV	618
QY	1018	DHLRREITERENQLTQSQKTMELKTTCTMLTEQVMDLEALNDELLEKEROWEAWRSVLG	1077
Db	619	DHLRREITERENQLTQSQKTMELKTTCTMLTEQVMDLEALNDELLEKEROWEAWRSVLG	678
QY	1078	DEKSOFEQCRVRELQMLDTEKOSRARADQRIITESQVVELAVKEHKABILALQALKEQK	1137
Db	679	DEKSOFEQCRVRELQMLDTEKOSRARADQRIITESQVVELAVKEHKABILALQALKEQK	738
QY	1138	LKAESLSKINDLEKXHAMLENNAESLOQKLETERELKORLLEEQAKLOQOQMDLQKNHIF	1197
Db	739	LKAESLSKINDLEKXHAMLENNAESLOQKLETERELKORLLEEQAKLOQOQMDLQKNHIF	798
QY	1198	RITQGLQALDRADLLKTERSLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKMD	1257
Db	799	RITQGLQALDRADLLKTERSLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKMD	858
QY	1258	QPAKKKKGLFSRRKEDPALPTQVPIQYNELKLALEKEKARCAEALQKTRIELRSARE	1317
Db	859	QPAKKKK-----VPIQYNELKLALEKEKARCAEALQKTRIELRSARE	903
QY	1318	EAAHKAHDHPSTPATARQOIAMSAIVRSEHOPQSAMSLLAPPSSRRKESSTPEFSR	1377
Db	904	EAAHKAHDHPSTPATARQOIAMSAIVRSEHOPQSAMSLLAPPSSRRKESSTPEFSR	963
QY	1378	RLKERMHNIHPRFNVGLNMRATKCAVCLDTHVFGQASKCLEQVCMCHPKCSTCLPATC	1437
Db	964	RLKERMHNIHPRFNVGLNMRATKCAVCLDTHVFGQASKCLEQVCMCHPKCSTCLPATC	1023
QY	1438	GLPAEYATHFTEAFCDKMNKSPGLQKTEPSSSLHLEGWMMKVRNNKRGQQQWDRKYIVLE	1497

RESULT 4

Q9QX19	ID	Q9QX19	PRELIMINARY;	PRT;	1618	AA.
AC	Q9QX19;					
DT	01-MAY-2000	(TrEMBLrel. 13, Created)				
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)				
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)				
DE	Postsynaptic density protein.					
GN	CITRON.					
OS	Rattus norvegicus (Rat).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
OX	NCBI TaxID=10116;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	Zhang W., Apperson M.L., Vasquez L.E., Kennedy M.B.;					
RT	"Citron, a PSD-95-binding protein at glutamatergic synapses on					
RT	inhibitory neurons."					
RL	Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.					
CC	-1- SIMILARITY: CONTAINS 1 PH DOMAIN.					
DR	EMBL: AF039218; AAC25483.1; --					
DR	GO: GO:0005822; C:intracellular; IEA.					
DR	GO: GO:0005324; F:ATP binding; IEA.					
DR	GO: GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.					
DR	GO: GO:0007242; P:intracellular signaling cascade; IEA.					
DR	InterPro: IPR001180; Citron.					
DR	InterPro: IPR005479; CPase L D2.					
DR	InterPro: IPR002219; DAG_P5-5ind.					
DR	InterPro: IPR001849; PH.					

```
DR InterPro; IPR000861; REM repeat.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG_PE_bind; 1.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR SMART; SM00109; C1; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS00867; CPASE_2; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR SEQUENCE 1618 AA; 185925 MW; EAED8079788AA459 CRC64;

Query Match 71.7%; Score 7521; DB 11; Length 1618;
Best Local Similarity 92.7%; Pred. No. 5.8e-308;
Matches 1500; Conservative 18; Mismatches 16; Indels 84; Gaps 5;

QY 468 MEQEMTLHRRVSEVPAVLSQKEVELKASGTORSLLLEQDLATYITECSLSKRSLEQARME 527
DB 1 MEQEMARLHRRVSEVPAVLSQKEVELKASGTORSLLLEQDLATYITECSLSKRSLEQARME 60

QY 528 VSQEDDKAQLLDHIREQSKQEIKEQYQAVQVEEMRLMMQLBEDLVASRRSDLYES 587
DB VSQEDDKAQLLDHIREQSKQEIKEQYQAVQVEEMRLMMQLBEDLVASRRSDLYES 120

QY 588 ELRESLAAEFPRKATECQKLLKAKDQKPEVGEYAKLEKINAEQQLKIQELQEKLEK 647
DB ELRESLAAEFPRKANECQKLLKAKDQKPEVGEYAKLEKINAEQQLKIQELQEKLEK 180

QY 648 A-----AKERAERLEKLNQEDSSEGIKKLVAEERHLENKVKRL 691
DB AKERAERLEKLNQIRQAKERAERLEKLNQEDSSEGIKKLVAEERHLENKVKRL 240

QY 692 ETMERENRLKDDIOTKSOIQOMADKILEBEKHEAQVSAQHLEVHLKQEQHYEKKI 751
DB ETMERENRLKDDIOTKSEIQOMADKILEBEKHEAQVSAQHLEVHLKQEQHYEKKI 300

QY 752 KYLDNQIKDLADKTELENNMQHHEBAHEKGIKILSEQKAMNAMDKIRSLERIVELS 811
DB KYLDNQIKDLADKESLETMMQHHEBAHEKGIKILSEQKAMNAMDKIRSLERIVELS 360

QY 812 EANKLAANSLFTQNNKAQEEIMSELROKFFLETQAGKLEAQNKLREQLKESHQDH 871
DB EANKLAANSLFTQNNKAQEEIMSELROKFFLETQAGKLEAQNKLREQLKESHQDH 420

QY 872 SDKNRLELETRLRVLSLEHEQKLEKQLTELQSLQERSQTLQAAPAALESQLR 931
DB SDKNRLELETRLRVLSLEHEQKLEKQLTELQSLQERSQTLQAAPAALESQLR 480

QY 932 QAKTELEETAEAEETIALTAHDEIQKFDALRNSCTVITDLBOLALQNTEDNAELNN 991
DB QAKTELEETAEAEETIALTAHDEIQKFDALRNSCTVITDLBOLALQNTEDNAELNN 540

QY 992 QNFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMOLTSQQTMEALKTTCTMLEEQ 1051
DB QNFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMOLTSQQTMEALKTTCTMLEEQ 600

QY 1052 VMDEALNDELLEKQEWAEVSLGDEKSOFCRVREIORMLDTEKQSRADORITIS 1111
DB VMDEALNDELLEKQEWAEVSLGDEKSOFCRVREIORMLDTEKQSRADORITIS 660

QY 1112 RQVVAVLAKHEKABEILALQALKEQKLAESLSDKLNLEKKHAMLENMARSLOQKLETE 1171
DB RQVVAVLAKHEKABEILALQALKEQKLAESLSDKLNLEKKHAMLENMARSLOQKLETE 720

QY 1172 RELKQRLLEEQAKLOQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYOLENIQVLY 1231
DB RELKQRLLEEQAKLOQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYOLENIQVLY 780

QY 1232 SHEKVWMEGTISQOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLAL 1291
DB SHEKVWMEGTISQOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLAL 840

QY 1292 EKEKARCABELEALQKTRIELRSAREEAAHAKATDHPHPSTPATARQQIAMSIVRSPSH 1351
DB EKEKARCABELEALQKTRIELRSAREEAAHAKATDHPHPSTPATARQQIAMSIVRSPSH 900

QY 1352 QPSAMSLAPSSRRKESSTPEPSRRLKERMHNNIPHRFNVLGNVNRATKCAVCLDTVHF 1411
DB QPSAMSLAPSSRRKESSTPEPSRRLKERMHNNIPHRFNVLGNVNRATKCAVCLDTVHF 960

QY 1412 GROASKCLEQVNMCHPKCSTCLPATCGLPAEYATHFEAFCDKNNSPGLQTKPESSSLH 1471
DB GROASKCLEQVNMCHPKCSTCLPATCGLPAEYATHFEAFCDKNNSPGLQTKPESSSLH 1020

QY 1472 LEGMKVPRNNKRGQGGWDRKIYVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGVSIIH 1531
DB LEGMKVPRNNKRGQGGWDRKIYVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGVSIIH 1080

QY 1532 GAYGASELANTAKA----- 1545
DB GAYGASELANTAKADVPYILKMESHPHTTCPORTIYLLAPSPDKQVWTALESVVAGG 1140

QY 1546 -----EKAADA-----KLGNSLLKLEGGDRLLDMNCTLPFSDQVVLVGTE 1586
DB -----EKAADA-----KLGNSLLKLEGGDRLLDMNCTLPFSDQVVLVGTE 1200

QY 1587 EGYALNVNLNSLTHVPGIGAVFQIYIIKDLKLLMAGEERALCLVDVKKVKQSLAQSH 1646
DB EGYALNVNLNSLTHVPGIGAVFQIYIIKDLKLLMAGEERALCLVDVKKVKQSLAQSH 1260

QY 1647 LPAQFDISPNIFEAVKCHLFGAGKIENGLCI CAAPSKVYVILRYNENLSKY CIRKEIET 1706
DB LPAQFDVSPNIFEAVKCHLFAAGKIENSLCI CAAPSKVYVILRYNENLSKY CIRKEIET 1320

QY 1707 SEPSCCHFTNYSLIGTNKFYIDMKQYTLBFLDKNDHSLAPAVFAASSNSFPYSIVQ 1766
DB SEPSCCHFTNYSLIGTNKFYIDMKQYTLBFLDKNDHSLAPAVFAASSNSFPYSIVQ 1380

QY 1767 VNSAQOREEYLLCFHEFGVFDVSGRRSRTDDLKWSRLPLAFAYREPFLVTFHNSLEVI 1826
DB VNSAQOREEYLLCFHEFGVFDVSGRRSRTDDLKWSRLPLAFAYREPFLVTFHNSLEVI 1440

QY 1827 EIQRSSAGTPARAYLDIPNRYLGPASSGAYLASSYQDKLVTCCKGNLYKESGTEH 1886
DB EIQRSSAGTPARAYLEIPNRYLGPASSGAYLASSYQDKLVTCCKGNLYKESGTEH 1500

QY 1887 HRGPSTSRSPNKGPPTYNEHITKRVASSPAPPEGSHHPREPSTPHRY--REGRTELRR 1944
DB HRGPSTSR--SPNKGPPTYNEHITKRVASSPAPPEGSHHPREPSTPHRYDREGRTELRR 1559

QY 1945 DKSPGRLERKSPGRMLSTRERSPQRLFEDSSRGLPAGAVRTPLSQVKNKRGQSA 2002
DB DKSPGRLERKSPGRMLSTRERSPQRLFEDSSRGLPAGAVRTPLSQVKNKWDQSS 1617

RESULT 5
Q8C1J1 PRELIMINARY; PRT; 1032 AA.
AC Q8C1J1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to citron (hypothetical protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=FVB/N; TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023775; AAH23775.1; -;
```

DR EMBL; BC051165.1; -.
 DR GO: 0005524; F:ATP binding; IEA.
 DR GO: 0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
 DR GO: 0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR005479; CPhase_L_D2.
 DR InterPro; IPR002219; DAG_PE_BIND.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00130; DAG_PE-bind; 1.
 DR Pfam; PF00169; PH; 1.
 DR SMART; SM00109; CL; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00233; PH; 1.
 DR PROSITE; PS00867; CPhase_2; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS00881; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS00003; PH_DOMAIN; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1032 AA; 116919 MW; 59AEDCF9659DA33B CRC64;
 Query Match 45.9%; Score 4819; DB 11; Length 1032;
 Best Local Similarity 91.5%; Pred. No. 1.2e-194;
 Matches 944; Conservative 10; Mismatches 10; Indels 68; Gaps 4;
 QY 1038 MEALTKTTCTMLEQVMDLEALNDELLEKERQWEAWRSVLGDEKSFQECRVRELQRLDTE 1097
 DB 1 MEALTKTTCTMLEQVMDLEALNDELLEKERQWEAWRSVLGDEKSFQECRVRELQRLDTE 60
 QY 1098 KOSRARADORITTESQVSLAYKEHKAETILALQOALKKEQKLAESLSDKLNLEKXHAML 1157
 DB 61 KOSRARADORITTESQVSLAYKEHKAETILALQOALKKEQKLAESLSDKLNLEKXHAML 120
 QY 1158 ENKARSLOQKLTETRELKORLLEEQAKLQOQMDLQKNHIFRITQGLQBALRADLLKTER 1217
 DB 121 ENKARSLOQKLTETRELKORLLEEQAKLQOQMDLQKNHIFRITQGLQBALRADLLKTER 180
 QY 1218 SDLEYOLENIQVLYSHEKVMGTTISQQTKLIDFLQAKVDQPAKKKKGLFSRRKEDPALP 1277
 DB 181 SDLEYOLENIQVLYSHEKVMGTTISQQTKLIDFLQAKVDQPAKKKKGLFSRRKEDPALP 240
 QY 1278 TVPQLOYNELKALKEKARCAELAEALOKTRIELRSAREEAHAKATDHPHSPATAPAR 1337
 DB 241 TVPQLOYNELKALKEKARCAELAEALOKTRIELRSAREEAHAKATDHPHSPATAPAR 300
 QY 1338 QIAMSALVRSPEHQSAMSLLAPPSSRRKESSTPEFSRRUKERMHNIPHRFNVGLNM 1397
 DB 301 QIAMSALVRSPEHQSAMSLLAPPSSRRKESSTPEFSRRUKERMHNIPHRFNVGLNM 360
 QY 1398 RATKCAVCLDTVHFGQASKCLEQVMCHPKCSTCLPATCGLPABYATHFTFAFCRDKN 1457
 DB 361 RATKCAVCLDTVHFGQASKCLEQVMCHPKCSTCLPATCGLPABYATHFTFAFCRDKN 420
 QY 1458 SPGLQTKESPSSSLHLEGWMMKVRNNKRGQGWDRKYIVLEGSKVLIYDNEAREAGQRPVE 1517
 DB 421 SPGLQTKESPSSSLHLEGWMMKVRNNKRGQGWDRKYIVLEGSKVLIYDNEAREAGQRPVE 480
 QY 1518 EFELCLPDGDVSIHGAVGASELAN TAKA----- 1545
 DB 481 EFELCLPDGDVSIHGAVGASELAN TAKA----- 540
 QY 1546 -----EKARADA----- 1572
 DB 541 QRWVTALESVAVAGVRVRKADARADITSEHQPVWVKLLGNSLLKLEGDDRLDMNC 600
 QY 1573 TLFPDQVVLVTGTEGLYALNVLKNSLTHVPVGIGAVFQYIYIKLEKLLMIAGEERALCL 1632
 DB 601 TLFPDQVVLVTGTEGLYALNVLKNSLTHVPVGIGAVFQYIYIKLEKLLMIAGEERALCL 660
 QY 1633 VDVKKVKQSLAOSHLPQAQPDIPSPNIFEA VKGCHLFGAGKIENGLICICAMPSKVILRYN 1692
 DB 661 VDVKKVKQSLAOSHLPQAQPDIPSPNIFEA VKGCHLFGAGKIENGLICICAMPSKVILRYN 720

QY 1693 ENLSKYCIRKEIETSEPCSCIHTFNYSILICTNKFYEDIMKQYTLSEFLDKNDHSLAPV 1752
 DB 721 DNLSKYCIRKEIETSEPCSCIHTFNYSILICTNKFYEDIMKQYTLSEFLDKNDHSLAPV 780
 QY 1753 FAASNSFPVSVIVQNSAGQREERYLLCFHEFGVFVDSYGRSRRTDDLKWSRLPLAFAYRE 1812
 DB 781 FAASNSFPVSVIVQNSAGQREERYLLCFHEFGVFVDSYGRSRRTDDLKWSRLPLAFAYRE 840
 QY 1813 PYLFVTHFNSLEVEICARSSAGTPARAYLDIPNRYLGPASISGAIYLASSYQDKLRVI 1872
 DB 841 PYLFVTHFNSLEVEICARSSAGTPARAYLDIPNRYLGPASISGAIYLASSYQDKLRVI 900
 QY 1873 CCKGNLVKESQTEHHRGPSTSRSSPNKRGPPPTNYEHITKRVASSPAPPEGSPHREPSTP 1932
 DB 901 CCKGNLVKESQTEHHRGPSTSRSSPNKRGPPPTNYEHITKRVASSPAPPEGSPHREPSTP 959
 QY 1933 HRY--REGRTELRDKSPGRPLEREKSPGRMLSTRRESRSPGRLEFSSRGRLPAGAVRTP 1990
 DB 960 HRYRDREGRTELRDKSPGRPLEREKSPGRMLSTRRESRSPGRLEFSSRGRLPAGAVRTP 1019
 QY 1991 LSQVKNKRGQSA 2002
 DB 1020 LSQVKNKWDQSS 1031
 RESULT 6
 O88937 PRELIMINARY; PRT; 494 AA.
 AC O88937;
 DT 01-NOV-1998 (TRENBUrel. 08, Created)
 DT 01-NOV-1998 (TRENBUrel. 08, last sequence update)
 DE 01-OCT-2003 (TRENBUrel. 25, last annotation update)
 DE Rho/rac-interacting citron kinase short isoform.
 GN CIT OR CRK-SK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99009084; PubMed=9792683;
 RA Di Cunto P., Calautti E., Hsiao J., Ong L., Topley G., Turco E.,
 RA Dotto G.P.;
 RT "Citron rho-interacting kinase, a novel tissue-specific Ser/Thr kinase
 RT encompassing the rho-rac-binding protein citron.";
 RL J. Biol. Chem. 273:29706-29711(1998).
 RC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF086823; AAC72822.1; -.
 DR HSSP; P00517; 1STC.
 DR MGI; MGI:105313; Cit.
 DR GO: 0005524; F:ATP binding; IEA.
 DR GO: 0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: 0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO: 0016740; F:transferase activity; IEA.
 DR GO: 0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000961; Kinase_C.
 DR InterPro; IPR007119; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase_C; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 494 AA; 55742 MW; 2C60D2294B28A185 CRC64;
 Query Match 20.6%; Score 2165.5; DB 11; Length 494;

Best Local Similarity 87.8%; Pred. No. 1.6e-83;
Matches 411; Conservative 26; Mismatches 30; Indels 1; Gaps 1;

QY 1 MLKFKYGARNPLDAGAAEPIASASRLNLFQFKPPFMTQQQMSPLSREGILDALFVLPF 60
DB 1 MLKFKYGVNPPSEASEPIASASRLNLFQFKPPLMTQQQMSALSREGMLDALFALPF 60

QY 61 ECSQPALMKIKHNSNFKYSDTIAELQELQPSAKOFEVRSVLCGCHFAEVQVREKATG 120
DB 61 ECSQPALMKIKHNSNFKYSDTIAELQELQPSAKOFEVRSVLCGCHFAEVQVREKATG 120

QY 121 DIYAMKMKKALLAQAEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVMEYQPGG 180
DB 121 DIYAMKMKKALLAQAEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVMEYQPGG 180

QY 181 DLSLLNRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIPENILVDRTGHKLVDF 240
DB 181 DLSLLNRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIPENILVDRTGHKLVDF 240

QY 241 GSAKAKNSKKNVAKLPICTPDYMAPEVLTVMNGDCKGTGGLCDKWSGVIAVEMVYGR 300
DB 241 GSAKAKNSKKNVAKLPICTPDYMAPEVLTVMNGDCKGTGGLCDKWSGVIAVEMVYGR 300

QY 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLCGQKRLKFEGLCCHPFF 360
DB 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLCGQKRLKFEGLCCHPFF 360

QY 361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWSSPQSLSPSGSGBELPFVGF 420
DB 361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWSSPQSLSPSGSGBELPFVGF 420

QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQDSQDKCHKM 468
DB 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQDSQDKCHKV 467

RESULT 7

ID O88527 PRELIMINARY; PRT; 448 AA.

AC O88527;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Citron-K kinase (Fragment).
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RP SEQUENCE FROM N.A.
RA Madaule P., Eda M., Watanabe N., Fujisawa K., Matsuoka T., Bito H.,
RA Ishizaki T., Narumiyu S.;
RT "Role of Citron kinase as a target of the small GTPase Rho in
RT cyto kinesin";
RL Nature 01-0-0(1998).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF070065; AAC27932.1; -;
DR HSP; P00517; 1YDR.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000961; Pkinase C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr pkinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR001245; Tyr pkinase.
DR Pfam; PF000649; Pkinase; 1.
DR Pfam; PF00433; Pkinase; C; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Tyrc kinase; 1.
DR SMART; SM00220; S_TKC; 1.

DR SMART; SM00133; S_TK X; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 448 448
SQ SEQUENCE 448 AA; 50650 MW; 548CA5A8AE4A6394 CRC64;

Query Match 20.5%; Score 2150.5; DB 11; Length 448;
Best Local Similarity 90.6%; Pred. No. 6.2e-83;
Matches 407; Conservative 21; Mismatches 20; Indels 1; Gaps 1;

QY 1 MLKFKYGARNPLDAGAAEPIASASRLNLFQFKPPFMTQQQMSPLSREGILDALFVLPF 60
DB 1 MLKFKYGVNPPSEASEPIASASRLNLFQFKPPLMTQQQMSALSREGMLDALFVLPF 60

QY 61 ECSQPALMKIKHNSNFKYSDTIAELQELQPSAKOFEVRSVLCGCHFAEVQVREKATG 120
DB 61 ECSQPALMKIKHNSNFKYSDTIAELQELQPSAKOFEVRSVLCGCHFAEVQVREKATG 120

QY 121 DIYAMKMKKALLAQAEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVMEYQPGG 180
DB 121 DIYAMKMKKALLAQAEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVMEYQPGG 180

QY 181 DLSLLNRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIPENILVDRTGHKLVDF 240
DB 181 DLSLLNRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIPENILVDRTGHKLVDF 240

QY 241 GSAKAKNSKKNVAKLPICTPDYMAPEVLTVMNGDCKGTGGLCDKWSGVIAVEMVYGR 300
DB 241 GSAKAKNSKKNVAKLPICTPDYMAPEVLTVMNGDCKGTGGLCDKWSGVIAVEMVYGR 300

QY 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLCGQKRLKFEGLCCHPFF 360
DB 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLCGQKRLKFEGLCCHPFF 360

QY 361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWSSPQSLSPSGSGBELPFVGF 420
DB 361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWSSPQSLSPSGSGBELPFVGF 420

QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEK 449
DB 421 YSKALGILGRSESVVSGLDSPAKTSSMEK 448

RESULT 8

ID O9VTY8 PRELIMINARY; PRT; 1854 AA.

AC O9VTY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE CG:0522 protein.
GN CG:0522.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Shandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hsatin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulyov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Ananidis P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Buesam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Earfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kamiker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RP Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RP FlyBase;
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AE003541; AAF49907.2; -;
 DR FlyBase; FBgn0036295; CG10522.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR001180; Citron.
 DR InterPro; IPR002219; DAG PE-bind.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR008271; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00219; TY_KC; 1.
 DR PROSITE; PS00081; DAG PE BIND DOM 2; 1.
 DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 1854 AA; 211324 MW; 93E5B69FE63DDE1D CRC64;
 Query Match 17.58; Score 1839.5; DB 5; Length 1854;
 Best Local Similarity 27.4%; Pred.No. 3.6e-65;
 Matches 552; Conservative 399; Mismatches 670; Indels 397; Gaps 71;
 QY 18 EPIASRASRLNLFPOGKP-----PFMTQOQMSPLS-----REG 50
 DB 6 EPISVRTARLNNLILGKAGVCAKPGASGSGIPASTRSIVPVSTTSAAVAEAI 65
 QY 51 ILDALFVLFECSPALMK-IKHSNVFVRKYSOTIAELOLQPSAKDFVRSVLGCHPA 109
 DB 66 LLDAFCLLYNECDKDTLKGRDRIAEFVNKFRPIIEETRLKRVNADFLKIGQGYFG 125
 QY 110 EVQVVRKATGDIYAMKMKKALLAGEQVSFEERNILSRSTSPWIPOLQAFODKNH 169
 DB 126 NVHLVVERQNDIYAMKKIKKSVTTSQ----VKERDINSINSEWLNQAFQNDN 181
 QY 170 LYLMEEQPGGDLSSLNRYEDQDENLQIYLAELILAVHSVHLVYHRDIKPNILV 229
 DB 182 LYLVMYMPGGDLSSLMSRH-GPFDLAFYLAELTVALHTLHEMGYVHRDIKPNILI 240
 QY 230 DRTGHILVDFGSAKNSKNWNKALPIGTDPYMAPEVLTVNMGD--GKTVGLDCDW 287
 DB 241 DRFGHIXLADFGNAALDRDGHVLSLSPVGTPTDYIAPELIQTISTYKLSKSMHDDVSCDYW 300
 QY 288 SVGVIAEMIXGRSPFAEGTSARTFNNIM-----NFQFLKFPDDPKVSSDFLDLIQSL 341
 DB 301 SMGIIGYELICETTFPHEDNVHETYSKILSHCBESHLKELISFPADLKSVNVRNLIESL 360
 QY 342 LCGQKERIKFGLCCHFPFSKIDWNNTNRPFPVPTLKSDDDTSNFDE-----PEKN 394
 DB 361 VTNFSKLSYERIKNHPFFSEIPWGSIRSQVPPITPTVRSDDDTNSFEDGIRHKTREQG 420
 QY 395 SWVSSPCQLSPSGFSGBELPFVGFYSKALGILGRSESVVSGLDSPAKTSMMEKLLIK 454
 DB 421 VAKSLTNNKSNDFSGKDLFFIGYSFVH-----MEKSAISATTEKLEQK 466
 QY 455 SKS-----IQDSQDKCHKMEQEMTLH---RRVSEVAVLQKVELKASQTSRLLQDL 507
 DB 467 LKELLQKLKTRNEBISMLKQDLRAQQSLRKTNDKNSQVADAKMEIK--KLQIIEKTM 524
 QY 508 ATYTECSLSKRSLE-QARMEVSQEDDKALQLLHDIREQSRKLEIK---EQEYQAOVEE 563
 DB 525 E-LTTCKTQIKTLQSSAKIDDEMWKKEATITOLLRLNPKYBEAKIASQRIEKQLAD 582
 QY 564 MRLMWNQLEEDLVARRRSDIYSELRESRLAABEFKRAKATECOH---KLKAKDQCKPE 620
 DB 583 KK-----QELASTLQKLDAREL-----EFNAKPECKHLSMKLQNYKDM--- 621
 QY 621 VGEVAKLEKINAEQOLKIQLQELQKLEAKKARERLEKLNQR-EDSEGRKKLVEARE 679

Db 622 -----LQIJK-EQNLKSETHBQRQMAELYEQLTDLRKKVRSQDNTNRTWEIKE 674
 QY 680 RHHSLEKNV-----KLETHBEREN-----RUKDIOYKSOQ 711
 Db 675 IRTELDSEISSKSTQBAXNATENIEIILRNEETIASNHELHAEKVLETKLQIKENE 734
 QY 712 IQWADKILEEKEHRAQVSAOHLVHLKQEHY-----EKIKVDNQIKKLADK 765
 Db 735 TQVRAECHRLERLQLAECRCQLAESSLATQVPYETAPGSLTELNAEDQLRADL--- 791
 QY 766 ETLENNQRHEEZAHEKGLISQKAMINAMWSKIRSLQRIIVELSEANKLAANSLEFQ 825
 Db 792 -----LAAKESNHQKGR-----ADQLTVTKLEQMLERFNEQSLSPKTS 832
 QY 826 RNKCAQEMISLRQOKFYLETQAGLEAONRKLEROLEKISHODHSDKN-----RL 878
 Db 833 HSRKQKQEG-----ETVGMLEQNEKLEKLAAREQMIQVQAARTANLSW 880
 QY 879 ELERLREV-----SLEHBEQKL-----ELKRLTELQLSQERESQITALQ- 920
 Db 881 KVEKQLEELSEKLLARMLTEDIKKYQNASDEAQMMLKTSQETQRESRIEELQ 940
 QY 921 -----APAALESQLRQAKTELETTAAEHEIQALTAHRDEIQKEDALRNSCTVITDLE 976
 Db 941 ELAAKRDVLKEHROWEKAEQRMCKSEIIEHLANVHLEQOE-----TEL 988
 QY 977 EQLNQLTE--DNAELANNQNFY--LSQKQDEASGANDEIVQLRSEVDHLRREITEREMOLT 1032
 Db 989 QKLRQIQSRDGVTLQKNTVIRELQERESKRKANDSCLVLQKEL-----KQIT 1037
 QY 1033 SQQWTEALKTCTMLBEQWMLBALNDELLEKEROWEAWSVLGDEKQFCRCRVRELQ 1092
 Db 1038 ---DNFORLYACISITDSQLTEVETLWKSEQERNSQKSLQDTHLEKLERNDQLTLEK 1094
 QY 1093 MLDEKQSRARADQ---ITESQVVELAVKEHKAELALQALKEQKLKAEBSLSDKND 1149
 Db 1095 QLTIVSEKELAQRAVLASEIDELRLNLKEQOKKLVAQQQLVEQ-----TNALFA 1147
 QY 1150 LEKKHAWLE-MNARSLOQKLETERELKORLLEBAQLOQMDLQKNHIFRL---TQGLQE 1205
 Db 1148 TQBRALLDQONANYEAQTADSNRENVLSKEENARILSELFHKKEEVENQLOAETRIGLES 1206
 QY 1206 ALBRADLLKTERSDELVENIQVLSHEKVKVKEGTISQTKLIDFLQAKM-DQPAKKK 1264
 Db 1207 A-QAN-LHAEIDSLQDTLAEKEQFTVQDRDKSNATLQKKLIDLYLQKVEDLSAKKK 1263
 QY 1265 GL-----FSRKKEDPALPTQV---PLQYNELKALEKAKAELEALQKTR--IELRS 1314
 Db 1264 TLADKLFGSSHTNKENVSPNDVSSILYRALKEELKREQKNSLLKEQLAQLNGTATLRS 1323
 QY 1315 AREEAHRKATDHPHSTATARQQTAMSAIVSRPEHOPSMSLLAPSSRRKESSTPEE 1374
 Db 1324 PRISAANGDSADP-----KORPVSIALPRSPQKQOQP---LKRTTSQVELKXTAK 1373
 QY 1375 FSRLEKXERHNIHPHFNVGL---NMRATKCAVCLDTHVFGQAQKLEQCMVCHPKGST 1431
 Db 1374 PTKVTIENQAH---HRFELAQESKYDANCVCVEKAVVAGSPFWKCKEVDTHRKCRS 1430
 QY 1432 CLPATCGLFAEYATHTEAFCDKQNSPGLQTKERSSLSHLLEGWKMVPRNNKQGGQWDR 1491
 Db 1431 NVQSHCGSTKTPAP-----SADDLSISQSVSSLTLDLS-VDVAGGTTSGGGE--- 1474
 QY 1492 KYIVLFGSKVLIYDNEAREAGQRPVEEFELCLPDGDSVSHGAVGASELANATAKAEKAD 1551
 Db 1475 -YI---GS-LVYSSDGAEE-----DQARKB----- 1493
 QY 1552 AKLGNLSLLKLEGGDLDMNCTLPFSD-QVILVGTTEGLYALNVLKNSLTHVPGIGAVFQ 1610
 Db 1494 -----IEYNCAFEVAEQIILLGCGNTGLYAVHLDQSRVLHISGLSVSC 1537
 QY 1611 IYIIXDLKLLMI--AGEBERALCLVDVKKV---OSLAQSHCLPAPQPDLS-----PNIEFA 1660
 Db 1538 MSICKRLAKAIWGTVGEKLYQC--DYQLERSCQSSSCHKFPVLETSAILPFPANRTPS 1595

QY 1661 VKGCHLFAGKIENGL--CICAMPSPKVVIILRYNENLSKYCIRKEIETSPSCSIHFTNY 1718
 Db 1596 EKWLVLISDEAENALDSVAJATSTRIVILKYLKLHMFKPVRLDTPATVTSIFPTRH 1655
 QY 1719 SLITGNTFYIDIMKOYTLTEFLDKNDHSLAPAFVAASSNSFPVSIQVNSAGOREEYLL 1778
 Db 1656 SAIVSSDFEYIDLDNYAAEEFVLSDKSM-----ESTAKQPLTAVRIS---RQEVLL 1706
 QY 1779 CFHEGCVDSYGRSRTDDLKWSRLPLAFAYRPPYLFVTHFNSLEVIEI----- 1828
 Db 1707 CFAEYGVFVDFBGCRSPYDLNMYAYPTGFVYRDPFLFISHYQSVQIVRLHRSRPSKEMAS 1766
 QY 1829 --QARSSAGTP--ARAYLDIENPRYLGPAI--SSGAIYL 1861
 Db 1767 GDNASESEFELQRYL-----PHYMSTLLANSQGVNL 1800
 RESULT 9
 OS4874
 ID O54874 PRELIMINARY; PRT; 1732 AA.
 AC O54874;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Myotonic dystrophy kinase-related Cdc42-binding kinase.
 DE MRCK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98078670; PubMed=9418861;
 RA Leung T., Chen X.Q., Tan L., Manser E., Lim L.;
 RT "Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a
 Cdc42 effector in promoting cytoskeletal reorganization.";
 RL Mol. Cell. Biol. 18:130-140(1998).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
 DR EMBL; AF021935; AAC02941.1; -.
 DR PIR; T14039; T14039.
 DR HSP; P28867; IFTQ.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR000095; PAKbox/Rho-binding.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000961; Pkinase C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00130; DAG_PE-bind; 1.
 DR Pfam; PF00786; PBD; 1.
 DR Pfam; PF00189; PH; 1.
 DR Pfam; PF00089; pkinase; 1.
 DR Pfam; PF00433; pkinase C; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00285; PBD; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00133; S_TK_X; 1.

DR PROSITE; P850108; CRIB; 1.
DR PROSITE; P800615; C TYPE LECTIN 1; 1.
DR PROSITE; P800479; DAG PE BIND DOM 1; 1.
DR PROSITE; P850081; DAG PE BIND DOM 2; 1.
DR PROSITE; P850003; PH DOMAIN; 1.
DR PROSITE; P800107; PROTEIN KINASE ATP; 1.
DR PROSITE; P850011; PROTEIN KINASE DOM; 1.
DR PROSITE; P850108; PROTEIN KINASE ST; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1732 AA; 197063 MW; 65F62F7872ACCD3B CRC64;

Query Match 13.8%; Score 1450; DB 11; Length 1732;
Best Local Similarity 24.6%; Pred. No. 7.8e-53;
Matches 521; Conservative 319; Mismatches 648; Indels 634; Gaps 75;

QY 36 PFWTQOQSPLEGGILDALFVFECSQPALMKIKHVSNFYKYSDDTAELQELQPSAK 95
DB 18 PAQTNGQC--FSVETLDDILICLYDECNPSLRREKNTILEYLEWAKPTSKVKQNRHRE 75
QY 96 DFEVSLVCGGHFAEVQVVRKATGDIYAMVKKALLAQOVQFFPEERNILSRSTSP 155
DB 76 DFEILLKVGAFGEVAVVVKLVADKVFAFKILNKWMLKRAETACFREERDVLVNGDSK 135
QY 156 WIPQOYAFQDKNHLVMBEYQPGDLSLNLNRYEDQDENLIQVLAELILAVHSHLM 215
DB 136 WITTLHAFQDDNNLYLVMDYTVGGDLLTLKSFEDRUPENARFYLAEMVTAIDSVHQL 195
QY 216 GYVHRDIPENILVDRTGHILKLVDFGSAKNNNSNMVNAKLPIDTPDYMAPEVLVWNGD 275
DB 196 HYVHRDIPKDNILMDMNGHILRADFGSKLMDGTQSSVAVGTPDIVISPEILQAMB-D 254
QY 276 GKGTYGLDCDWSVGVYAIEMLYIGRSPFAGTSARTFNIMKFORFLKPPDD-PKVSDSF 334
DB 255 GGRGYPCEDWWSLGVCMYEMLYGETPFYAESLVETYGKIMNHKRFQPTQVTDVSENA 314
QY 335 LDLIOSLLCGOKERLKFEG--CHPPFSKDWNNIRNSPPFPVPTLKSDDDTSNFDPEP 391
DB 315 KDLIRRLICSRHRLGONGIEFKHPFFSGIDWNNIRNCEAPY:PEVSSPTDTSNFDVD 374
QY 392 E---KNSWSSPQCQLSPGSGEELPFVFGYSKALGILGS-----ESV 434
DB 375 DDCLNSETMPPP--THTAFSGHLPFVFGTYTSTCVLSRSLCLRTVAGTSLDLVDV 431
QY 435 VSGDSPAKTSMWEKLLIKSKELQDSQDKHMQEMLRHRYSB---VEA-----484
DB 432 QRTLNNNLATEAYERKI-----KRLQEKLELRKLQESTQTVOALQYSTV 477
QY 485 ---VLSQKEVELKASGTQSLLEQDLATYITECSSLRSLRLEQARVEVSQEDDKALQLLHD 541
DB 478 DGLPTASKDLKSLKEETELRKQVA---EVNHLQQLLEBAN-SVARELDDAFQIKA 532
QY 542 IREQSRKLOEIKEOYQAOQVEEMRLMMNQLEEDLYSARRRSDLYSESLAESLAAEFPR 601
DB 533 FEKQIKTLQOERE-----ELNKELVQASERLNQSKELKDA-----568
QY 602 KATECHKLLKAKDQGPVGVYAKLEKINAEQQLKIOELQEKLEAKAKERAERELEKIQ 661
DB 569 ---HCQKRL-----ANQEFMEINE-----RLTELHTQKQKLAHVYDKB-EVVD 608
QY 662 NREDSSEGRKXKLVAEERHSLKNVKRLETWERENRLKDDIOTKQOIQOQMAKXILE 721
DB 609 LVNQKAESLRQELRRAERAKKELEVHTEALIAEASKDKRLREQSRHYSKQLEN-----662
QY 722 LBEKHREAQVSAQHLEVHLKQEQHYEKKIVLDNQ-----LKKDLADKETLNNMQRHE 776
DB 663 LE-----GLKQKQISYSPGICSEHQEITKUTDLEKKSIF-----YE 701
QY 777 EEAHEXGKILSEQKAMINAMDSKIRLEQRIYVELSEANKLAANSLSFTQRNNKQAEEMIS 836
DB 702 EBISKREGIHA-----SEIKNLKXELHD-SEGGQALANKSEIMV-----738
QY 837 ELRQKFXYLETQAGKLEAQRNKLBEQLEKISHQDSDDKNRLLLELTRLEVSLEHEBOKL 896

DB 739 -----LKDLEK-----TR-RESQSEREEFEN 759
QY 897 ELKQLTELQLSQERESQLTALQAARAALESQLRQAkteLEETTAEEAEIQAHTARD 956
DB 760 EFKQ-----YEREKVL-----771
QY 957 EIQRKFDALRNSCTVITDLSEQLNQLTEDNAELNNQNFYLSKOLDEASGANDEIVQRSE 1016
DB 772 -----LTEENKULT-----SE 782
QY 1017 VDHLRREITEREMQLTQKQTMELKTTCTYMLEEQVMDLEALNDELLEKEREQWEAERSVL 1076
DB 783 LD-----KLTs-----LYESLSLRNQHLEEEVKDLA-----DKKESVAHWEA-----819
QY 1077 GDEKSQFEQVRRELQRMDETEKQSRFAADQORITESQVVELAVKHEKAEIALLAQALKEQ 1136
DB 820 -----QITEIIQWVSDEKDARG-----YLOALASQMTETE 850
QY 1137 KLKAEISLSDKINDLE---KHAMLENNAR-SLOOKLETERELKQRLLEBOAKLQQQMDLQ 1192
DB 851 ALRNSSLGTRATDMPWQWRFPKIDMSARLELSALDAEIRAQKAIQEEELNKVK-----904
QY 1193 KWHIFRLTQGLQBALDRADLLKTERSDELYOLENIQVLSHEKVKMEGTISQOTKLIDFL 1252
DB 905 -----ASNITECKLKDEKKNLELLSELEQLIKTEBLRSEKGVHRDSQHSFL 954
QY 1253 QAKVDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALEKAKARCAELBEALQKRIEL 1312
DB 955 AFLNTP-----DALDQPE---968
QY 1313 RSAREEAAHRKATDHPHPS--TPATARQOIAVSAIVRSEPHOPSAMSLAPSPSRKESST 1371
DB 969 RS-----PSTPAGKGRRIADAPL--PVHTPTL-----RKKGCPA 1002
QY 1372 PEEFSRLKERMHHNIPIHRFNVGLNMEATKCAVCLD-TVHFGROASKLEQCVNCHPKCS 1430
DB 1003 SAGFPKPKK-----THQFVKSFAPTCKHOCTSLMVGLIRQGCSECVGFSCHITCV 1055
QY 1431 TCLPATCGLPAEYATHTEAFRCRDKKNSPGLQKEP-----SSSLHLEGMKVPNNK 1483
DB 1056 NKAFTTCFVPE-----QTKPLGIDPQKGVGAYEGHVRIPK-P 1094
QY 1484 RG-QOQWDRKVIIVLEGSKVLIYDNEAREAGORPV-----EEPELCLPDGDVSIHG 1532
DB 1095 AGVKKGHQRALAVVCDFKLFLDYIAEGKASQSPSSVISQVIDMRDEEFSSVSLASDVIIA 1154
QY 1533 A-----VGASELANTAKAEK-----ADAKL-----IGNSLCLKEGDRL 1568
DB 1155 SRKDIPICIFRVTSQ-----SAPSOKCSILMLADSETERSKMWGVLSLHVKLKNKFRDRS 1212
QY 1569 -----DMNCTLP-----FSDQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFQIY 1612
DB 1213 VYVPKEAYDSTPLIKTQAAAIIDHERVALGNEEGLFVVHTKDEIRVGDNKKKHQIE 1272
QY 1613 IIKDELKLLIAGEERALCLVDVKVKQSLAQSHLPQAQDISPNIFEAVKCHLFGAGKI 1672
DB 1273 LIPSDQLVAVISGRNHRVLPFPMALDGRETFVYKLA-----TKGQOTIAAGKV 1322
QY 1673 ENGL--CICAMPKSVILRYNNENLSKYCRKEIETSEPCSCIHFTNYS--ILIGTNKFY 1728
DB 1323 RHGALSCLCVAMKQVLCYELFQSKTRHKFKIQV--PCNVQMAIFSEHLVCG----F 1376
QY 1729 BIDMKQVYLE-----EFLDNHSLAPAVFAASSNSFPVSIQVNSAGOREEYLLCFHE 1782
DB 1377 QSGFLRPLNAGEGSPCNMLHSHNDHTLA--FITHQPMDAICAVEISN---KEYLLCFSS 1429
QY 1783 FGVPVDSYGRSRRTDCLKWSRLPLAFAYREPYLIVTHFNLSLEIVIEIQARSAGTAPARAYL 1842
DB 1430 IGIYTDCCQRRSRQOELMWPANFPSSCCYNAPYLSIYSENNAVDFIDVNSMEWIQT-----L 1484
QY 1843 DIPNPRYLGAISGGAITYLASSYQDKLRVICCKGNLVKES-----GTEHHRGPSTSPSS 1897
DB 1485 PLKKVR---PUNTEGSUNLIG--LETIRLILYFKMKMAEGDELVVPETSDNSRKQWNRNIN 1539

SQ SEQUENCE 1760 AA; 199209 MW; 3A1CA9A7A4B5FA3 CRC64;
Query Match 13.2%; Score 1385.5; DB 4; Length 1760;
Best Local Similarity 24.6%; Pred. No. 4.1e-50;
Matches 513; Conservative 324; Mismatches 642; Indels 607; Gaps 80;
QY 44 SPISREGILDALFVLEECGAPALMIXHVSFVRKYSDTIAELOELQPSAKDFEVSIV 103
DB 72 SAUSVETLLDVCLYTECHSALRDKYVAEFLEWAKFTQLVKEMQHUHREDFIIVKI 131
QY 104 GCCHFAEVOVREKATCDIYAMVKKKALLAQEQVSPFEEBENIISRSTSPWLOQVA 163
DB 132 GRGAFGEVAVVMKNTIRIYAMKILNKWEMLEKAEATACFREEDVLVNGDCQWITALHYA 191
QY 164 FQDNHLYLMEEYQPGDILLNRYEDQIDENLOFYLAELILAVSHVLMGYVHRDIK 223
DB 192 FODENHLYLVMDYVGGDILLTLKSFEDKLPEDMARFYIGEMVLATDSHQHLYVHRDIK 251
QY 224 PENILVDRTHIKLVDPFGSAAKXNSKMNNAKLPIGTDPYMAPEVLTVNMGDKGTGYLD 283
DB 252 PDNVLVDNGHIRLADFGSLKXNDGTVQSSVAVGTPTYISPEILQAME-DGMKYGPE 310
QY 284 CDWWSGVVIAYEMIYGRSPFAETSARTNNINMFORFLKPPD-PKVSSDFLDLIQSLL 342
DB 311 CDWWSLGVCMYEMLYGETPFYAESLVEITYKIMNHEERFPFPHVTDVSEEAADLIORLI 370
QY 343 CGOKERLKPTEL--CCHPFPSKIDWNTRNSPPFPVPTLKSDDDTSNPFDEPE--KNSW 396
DB 371 CSHERLQNGIEDFKHAFEGELNWNENRLEAPYIPDVSSPDSNFDVDDVLRNTE 430
QY 397 VSSSPQOLSP--SGFSGBELPFGVFSYKALGILGRSESVVSGDSPAKTSSMEKLLI 453
DB 431 I-----LPPSGHTGFGSLHLPFIGTFT-----TESCF--DRGSLKSIQMSNTLT 474
QY 454 KSELOQSDQKCHKMOETRLHRRVSEVAVLSQKEVELKASETQRSLLLEQDLATYITE 513
DB 475 KDEDVQ--RDLESLQME--AVERRI----- 496
QY 514 CSSLKSLEQARMEVSQEDKALQLLHDIREOSRKLQEI-KEQEQYQAQVEMRLMNOLE 572
DB 497 -----RELQKLELSRKLQESQTQSLHSGSRALSNSNRDKIKLNEEIERLKNKIA 551
QY 573 EDIVSARRSDLYESIELRSRLAAEFKFKATECQHKLLKAKDQCKPEVGEYAKLEKINA 632
DB 552 D-----SNRLERQL-EDTVALQ-----ERDS 573
QY 633 EQQKIQEIQEKLKAAKERAERELEKLNREDSSEGIRKLVAEERHRSLENKVKLE 692
DB 574 TQRLGLEQKQHVVRQKEE-----LHQLVSEASER----- 604
QY 693 TWERRENRLKDDITQTSQOIQQWADKILEBEKREHAQVSAQHLEVLHQKQKHVEEKIK 752
DB 605 -LKSQAKELKDAHQKALQEQFS---ELNRMELRAQCKVSRQLRDKEEEMEVATQ 659
QY 753 VLDNQIKKDLAKETLENMORHEBEAHEKGLSPQKAWINAMSKIRSLEQRTVELSE 812
DB 660 KVD-----AMQENRAB-----KLKLELEA-----QLDDAAVEASK 691
QY 813 ANKLAANSILFTORNKAQEMISELRQOKFYLETQAGKLEAQNRKLEBQLEKISHQDHS 872
DB 692 ERKLREHSENF-----KQMESELEALKV--KQGRGAGATLEHQEQEISKIS---- 737
QY 873 DKORLLETRLEVSLEHEEQLEKRLQTLQLSLORESQLTALQAARALESQLRQ 932
DB 738 -----ELEKKV--LFVEE---ELVRREASHVLEVRKVKVEHDSHQHALQKELM 784
QY 933 AKTELETTAAEBEETIQAHTADEIQRKFDALNRSCVTITDLEBQNLQTLTDENAELNNQ 992
DB 785 LKDKLEKSKRERHNEEAVG-----TIKDYERERAMLFDENKKTAE 828
QY 993 NFVLSQLODEASGANDIEIQLARSEVDHLRREITEREMQLTQKQWTEALKTTCTMLEBQV 1052
DB 829 NEXLCSFVD-----KLTQNR-----QLEDEL 850

QY 1053 MDLEALNDELLEKEROHEAWRSVLGDEKSFQCEVRBELQMLDTEKQSRARADQRTESR 1112
DB 851 QDLAALK-----ESVAHWEA-----QIAEIIQWVSDEKD--ARGYLQALASK 890
QY 1113 QVVELAVKHAETLALQOALKEQKLKAESLSKLDNLEKKHAMLENMAR-SLQOKLETE 1171
DB 891 MTEL-----EALRSSLSGRTL-DPLMKV-RRSQKLDMSARLEQLSALEAE 935
QY 1172 RELKORLLEQALQO-QMDLQKNHIFRLTOGQ-LEALDRADILLKTERSULEYLEJENIQL 1230
DB 936 IRAKOLVQEBELRVKQDANLTLSEK-----LKDSEAKNRELLEEMEL 977
QY 1231 YSHEKVMKEGTISQQT--KLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQVNLK 1288
DB 978 ----KKQMEKFRADTGLKLPDF-----QDSIFE-----YENTAP 1008
QY 1289 LALEKERCARCAELEALQKTRIELRSAREEAAHRKATDHPHS-----TPATARQOIAM 1343
DB 1009 LAHD-----LTFRTSSASEQETQAPKPEASPSMSVAASEQOEDMA 1048
QY 1344 AIVRSPHQPSAMSLAPPSSRRKESSTPEFSRRLKERVHHNI PHRENVGLNMRATKA 1403
DB 1049 ----RPPQRESAVPL---PTQALALAGPKPA-----HQFSIKSFSTQCS 1089
QY 1404 VCLD-TVHFRQASKCLEQCMCHPKCSTCLPATCGLPAEYATHFTEAFCDKXNSPGLQ 1462
DB 1090 HCTSLMVLIRQGVACBVCFSFACHVSCKDGAQVCPPIPE-----Q 1130
QY 1463 TKEP-----SSSLHLEGWKMUPRNKRQOQCKQDRKIYVLEGSKVLIVD-NEAREAGQR 1514
DB 1131 SKRPLGVDVORGITAYKGVHKVPKPT-GVKKQWQRAYAVCDCKFLYDLDPGKSTQPG 1189
QY 1515 PVBEFELCLPDGVSIGHAVGASELANATAKAEA--EADAKLLG-----NSLLKLEGGD 1566
DB 1190 VIASQVLDLRDEFSV-SSVLASDVITHATRPDIPCFRVTASILGAPSKTSSLLITENE 1248
QY 1567 RLKNNCTLPSSDQVVLVGTBEGYAL--NVLKNSLTHV-----PGIGAVFOIYI 1614
DB 1249 -----NEKKWVGILEGLQSLHKNRLNQVHVHVEAYDSPLIKAILTAI 1298
QY 1615 KLEKLMIAGEERALCLVDV-----KKVQK-SLA-----QSHLPAQ 1650
DB 1299 -DADRIAV--GLEGLYVIEVTRDVIVRAADCKVHQIELAPREKIVILLCGENHHVLY 1355
QY 1651 P-----DISPNIFE--AVKGCHLFGAGKIENGLCTCAAMPSKVILRYNNENLSKYCIK 1702
DB 1356 PWSSLDCAEGSFIDKLPETKGCQLMATATLKRNSGTCLFVAVKRLILCYEIQRTKPFHRK 1415
QY 1703 EIETSEPCS--CIHFTNYSILITGNKFEYIDMQYILE-----EFLDKNDHSLAPVA 1754
DB 1416 FNEIVAPGSVOCLAVURDLRCVG-----YPSGFCLLSIGDQQLNVLNVPDLSLA--FL 1468
QY 1755 ASSNSFPVSVQVNSAGOREEYILLCFHEFGVFDYSYGRSRKTDLLKWSRLPLAFA----- 1809
DB 1469 SQQSFDCALCAVELES---BEYLLCFSHMGLYVDPQGRARARAOELMWPAPVACSCSPH 1524
QY 1810 ---YREPYLFVTHFNSLEVIE---IQARSAGT-----PAR-----AYLD 1843
DB 1525 VTVYSEYGVDFVDRTWEMVQTLGLRIRPLNSEGLNLNLCPEPPLIYFKSFSGAVLN 1584
QY 1844 IPN-----RYL-----GPAISSGAIVYASSYQ----- 1866
DB 1585 VPDTSDNSKKOMLTRSKRFRVFKVPEERLQORREMLRDPPELRKXNISNPTNFNVAHM 1644
QY 1867 ---DKLVI CCKGNLVKESGTEHHRGPSTSRSPKRGPPTYNE----- 1907
DB 1645 GPGDMQVMDLPLSAVPPSQEERPGPAPTNLA---RQPPSRNKPYISWPSGSGSEFSVT 1701
QY 1908 -----HITKRVASSPAPPEGPGHPRPSTPHR 1934
DB 1702 VPLRMSDDPDQDFDKPDSSTKH--STFSSNSNPSGPPSPNSPHR 1745

143	QY	FODENHLYVMVYVGGDLTLISKFDKLPEDMARFYIGEMVLAIDS THQLHVYHRDIK	202
164	QY	POCKNHYLYMEYQPGDGLSLNRYBDOJDENLIOFYLAELITLAVHSLMGVYHRDIK	222
224	QY	PENILVDRTHGHIKLAVDFGSAKNSNKNVNAKLPIGTPDYMAPEVLTVANGDGKGTGLD	283
203	Db	PONVLDDVNGHRIADFGCLKNDGDTVOSSVAVGTPDYISPEILQAME--DMGKGYBE	261
284	QY	CDWMSVGVIAYEMTYGRSPFABGTSARTFNNIMNFORFLKFPDD-PKVSDFDLDTQSLL	342
262	Db	CDWMSLGVCMYEMLYGTPTFVABSLVETYGKIMHBEERFOFQSHVTDVSEEAKDLIQRLL	321
343	QY	CGOKERLKFEGL---CCHPTFSKIDNNNIANSPPVPVTLKSDDDTSNDEFE---KNSW	396
322	Db	CSRERRLQONGIDFKCHAFEGUNWENIARLEAPIYDPVSPSDTSNFDVDDVLIRNTE	381
397	QY	VSSPFCQULSP---SGFSGEELPFGVSYSKALGILGRSESVVSGLDSPSAKTSMBEKKLLI	453
382	Db	I-----LPPGSHGTGFSGLHLPFTIGTFT-----TESCPS--DRGSLKIMQSNLTLT	425
454	QY	KSKELOSDQKCHWEGEMTELRHRRVSEVAVLSQKEVELKASETQORSILLEODLATYITE	513
426	Db	KDEDVQ---RDLEHSLQWE--AYERRI-----	447
514	QY	CSLSKRSLEQARMEVSGEDDKALQLLHDIREOSRKLQEI-KQEYQAOQVBEEMRLMMNOL	572
448	Db	-----RRLEQEKLELSKLOQESTOTVQSLHSGSRALSNSNRDKIKLNEETIERLKNKIA	502
573	QY	EDLYSARRRSDLYESELRESRLAABFFKRYKATCEQHKLKAKQDQKPEVGEYAKLEKINA	632
503	Db	D-----SNRLERQL-EDTVLRQ-----	524
633	QY	BQOLKIOBQEKLEKAKERAERELEKLQNRBDSSEGIKVLVEABERRHSLDENVKRLE	692
525	Db	TQRLRGLEKQHRVVRQKEE-----LHKQLVEASER-----	555
693	QY	TWEXRENRLKDDITQKSOQIQOMADKILBLEKXHREAQVSAQHLVHLKQKQHYEBEKIK	752
556	Db	-----LKSOAKELKHDAQOORKLALQFES-----ELNERNVIELRAQOKVSRQRJDRKEEEMEVATQ	610
753	QY	VLDNQIKDLADKETLENMWRREEBAHBKGKILSEQAMINAMDSKIRSLRQFVVELSE	812
611	Db	KVD-----AMRQEMRRAE-----KLRKELEA-----QLDDVAVAEASK	642
813	QY	ANKLAANSSLFTQENMKQAQEMSELROOKFVLETOAGKLEAQNKLBEQLEKISHQDHS	872
643	Db	ERKURESEHFC-----KQMESELEALKV---XQGGGAGATLHQOEISKTS-----	688
873	QY	DKNRLLLETRLREVSLEHEEQLEKLRQUTELQLSLOFRESQLTALQAAARAALESQLRQ	932
689	Db	-----ELEKVV---LFYEE---ELVREASHVLEKVNKVEHDSHQALQKEILM	735
933	QY	AKTELESTTAAEEETCALTAHDELTQRFDAIRNSCIVITDLESQLNQLTDEDAELNNO	992
736	Db	LKOKLEKSKERNHEEAVG-----TIKQYERERAMLFDENKKLTAE	779
993	QY	NFYLSKOLDEASGANDEIVQLRSEVDHLRREITEREMOLITSQOTMEALKTTCTMLEBOV	1052
780	Db	NEKLCSFVD-----XLTQNR-----QLEDEL	801
1053	QY	MDLEALNDELLEKROQHEAWRSVJGDEKQFECRVRELEORMLDTEKQSRADQRITES	1112
802	Db	QDLAAKK-----ESVAHWEA-----QIAEIIQWVSDEKD--ARGYIQLAASK	841
1113	QY	QVVELAVYKHKABILAQALKEQKLKABSLDKINDLEKHKHMLENNAR-SLOQKLETE	1171
842	Db	MTEEL-----BALRSSLSGRTL-DPLWKV-RRSQKLDMSARLELOSALEAE	886
1172	QY	RELKQRIILBERCAKLOQ-QMDLOKNHIFRLTQGLQALDRADLLKTERSDBLYQLENTQVL	1230
887	Db	IRAKOLQOELRVKQDANUTLESK-----LKDSEAKNRELLEERMEIL	928
1231	QY	YSHEKVXWECTISQOQT--KLIDFLOAKMDQPAKKKKGLFSRKEDPALPTQVPOVLYNELK	1288

Db 929 -----KKMBEKFADTGLKLPDF-----QDSIFE-----VFNTAP 959
 Qy 1289 LALEKEKARCAEAEALQKTRIELRSGARBEAAHRKATDHPHS-----TPATAROQIAMS 1343
 Db 960 LAHD-----LIFRTSSASEQETQAPKEASPSMSVAASEQEDMA 999
 Qy 1344 AIVRSEHQBSAMSLAPSSRRKESSTPEBSRRILKRMHINIHRFNVGLNMRATKA 1403
 Db 1000 -----RPPQRESAVPL-----PTQALVLGPKPKA-----HQFSIKFSFSTQCS 1040
 Qy 1404 VCLD-TVHFCRQASKLEQCVQCHPKCSTCLPATCGLPAEYATHFTAFCDKMNSPGLQ 1462
 Db 1041 HCTSLMWGLIRQYACEVCSFACHVCKGAPQVCPDPE-----Q 1081
 Qy 1463 TKEP-----SSSLHLEGMMKVPNNKRGQOQWDRKYYIVLEGSKVLIYD-NEAREAGOR 1514
 Db 1082 SKRPLGVDVQRGIGTAYKGHVKEPT-GVKKQWQRAYAVVCECKFLYDLPEGKSTQPG 1140
 Qy 1515 PVSEFELCLPDGDSVTHGAVGASELANATAEKA-----EADAKLLG-----NSLLKLEGGD 1566
 Db 1141 VIASQVLDLDRDBFSV-SSVLASDVHATRDIPCFIRVTASLLGAPSKTSSLLILTENE 1199
 Qy 1567 RLDNMCTLPSPDQVVLVTGEBEGYAL-----NVLKNSLTHV-----PGIGAVFOYII 1614
 Db 1200 -----NEKRWGILEGQSILHKURLNQVHVPLEAYDSSLPLIKAILTAIV 1249
 Qy 1615 KDLKLLMIGERALLVDV-----KVKQ-SLA-----QSHLPAQ 1650
 Db 1250 -DADTAV--GLBEGLYIVETRDVIRVADCKKHVHQIELAPREKIVILLCGRNHHVHLY 1306
 Qy 1651 P-----DISPNIPE-----AVKGCHLFGAGKIENGLCTCAAMPSKVWILRYNENLSKYCIK 1702
 Db 1307 PWSLDGAGSFIDIKLPETKGCQMATATLKRNSGTCLFVAVKRLILCYEIQRTKPFHRK 1366
 Qy 1703 EITSEPCS--CHFTNYSILITGNKFEYIDMKQYILE-----BFLDKDHS LAPAFA 1754
 Db 1367 FNBIVAPGSVQCLAVLRDLVCG-----YPSGFCLLSIGDGGQPLNIVNENPFLA---FL 1419
 Qy 1755 ASSNSRPVSVQVNSAGOREYLLCFHEFGVFDVSYGRSRTDDLKWSRLPLAFA----- 1809
 Db 1420 SQSFDALCAVELES-----BEYLLCFSHGGLYVDPQGRARAGELMWPAAPVACSCSPH 1475
 Qy 1810 ---YREPLYVTHFNLSLEVIE-----IQARSSAGT-----PAR-----AYLD 1843
 Db 1476 VTVYSEYGVDFVDTVMWQTTIGLRIRPLNSEGTLLNLNCEPPLLIYFKSKFGAVLN 1535
 Qy 1844 INP-----RYL-----GPAISSGAIYVASSYQ----- 1866
 Db 1536 VPDTSDNSKKQMLRTSKRNFVFKVPEERELQORREMLRDLPELRSKMSINPTNFNVHVM 1595
 Qy 1867 ---DKLRVICCKGNLYKESGTEHRRGPTSSRSPNKRGPPTVNE----- 1907
 Db 1596 GPGDMQVMDLPLSAVPPSQEERPGPATNLA---RQPPSRNKPYSWPSGGSPPSVT 1652
 Qy 1908 -----HITRVASSPAPPGPSHPREPSTPHR 1934
 Db 1653 VPLRSMDDPDQDFKEPDSSTKH---STPSNSNPNPSPSPNSPHR 1696

RESULT 13

Q7TT49 PRELIMINARY; PRT; 1713 AA.
 ID Q7TT49
 AC Q7TT49
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CDC42-binding protein kinase beta.
 GN CDC42BPB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 IX NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Huang C.Q., Wu S.L., Cheng Z.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY27590; AAF34403.1; -;
 KW Kinase.
 SQ SEQUENCE 1713 AA; 194886 MW; 80C999262C96DAA6 CRC64;
 Query March 13.18; Score 1373; DB 11; Length 1713;
 Best Local Similarity 24.78; Pred.No.1.3e-49;
 Matches 484; Conservative 306; Mismatches 625; Indels 546; Gaps 67;
 Qy 44 SPLREGILDALFVLFEECSQPALMKIKHVSFNVRKYSDTIAELOELQPSAKDFFVRSVLV 103
 Db 23 SLSVETLLDVLVCLYTECSHSALRRDKYVAEFLEWAKPFTQLVKDQLHREDFEIKVI 82
 Qy 104 GCGHFAEVQVVRKATGDIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPITPOLQYA 163
 Db 83 GRGAFGEVAVVFMKNTERIYAMKILNKWMLKRAETACFFREERDVLVNGDCQMITALHYA 142
 Qy 164 FODKNHLYLMBEYQPGDGLLSLLNRYEDQDENLIQYLAELILAVHSHVLMGYVHRDIK 223
 Db 143 FQDENYLYLVNDYVVGDLTLTLKFKEDLPEDKARFYIGENVLAIDSIHQLHVHRDIK 202
 Qy 224 PENILVDRTHIKLVDFGSAKKNKVNNAKLPIDGPIYMAPEVLTVMGDKGKGYGLD 283
 Db 203 PDNVLLDVNGHIRLADFGSLCKMNDGTQVQSSVAVGTPDYISPEILOAME-DGMGKYGPE 261
 Qy 284 CDWMSGVIAVEMLYGRSPRAEGTSARTENNINNFQRLKFPDD-PKVSSDFDLDLQSL 342
 Db 262 CDWWSLGVCMFEMLYGTPTFAESLVETYGKIMHEERFQFPFPHVTVSEAKQLQRLI 321
 Qy 343 CGQKERLKFEGL---CCHPFFSKIDMNNIRNSPPFVPTLTKSDDDTSNFDPE---KNSW 396
 Db 322 CSRERRLGQNGIEDFKKHAFFEGNLNENIRLEAPYIPDVSSPDSPTSNFVDDVLRNIE 381
 Qy 397 VSSSPCQLSP---SGFSGBELPFVGFYSKALGILGRSESVVGLSDSPAKTSMEKLLI 453
 Db 382 I-----LPPGSHTFGSLHLPGFTFT-----TESCFS---DRGSKSMIQSNLT 425
 Qy 454 KSKELQSDQKCHWQEMTLHRVSEVEAVLSQKEVELKASSETQSLLEQDLATYITE 513
 Db 426 KDEVDQDLENSLQIEAVERRIE----- 449
 Qy 514 CSSLKRSLQARMVS---QEDDKALQLLHDIRE---QSRKLEIKQEYQAOVVEEMLM 567
 Db 450 -----LEQKLESLKQLQESTQTQVSLHGSTRALGNSNRDKEIK--RLNEELERMKSK 500
 Qy 568 M---NOLE---EDLVARRRSDLYSELRESRLAAEEFKRKATECQHKLLKAKDQGRPEV 621
 Db 501 MADSNRLERQLEDVTLRQ----- 519
 Qy 622 GEYAKLEKINAEQQLKTOELQEKLEKAAKRAERELKLNQREDSSGIRKLVAEERR 681
 Db 520 -----EHEDSTQRLKGLKQYRLARQKEE-----LHKQLVEASER- 555
 Qy 682 HSLNKVKLETMERENRLKDDIQTSQOIQQWADKILEEKHREAQVSAQHLVHLK 741
 Db 556 -----LKSQTKELKDAHQQRKALQEPS-----ELNEMALRSQKQVXSQRLR 599
 Qy 742 QKEQHYBEKIKVLDNQIKKOLADKETLENMWMQREHEEAHEKGLSEKAMINAMDSEKIR 801
 Db 600 DKEEMEEMVAMQKIDS-MEQDIRKSE---KSRKELEAR----- 632
 Qy 802 SLEGRIVELSEANKLANSSILFTORNMKAQEMISELQKQFYLETQAGLEAQNRLKEE 861
 Db 633 -LEDAVAEASKERKLREHSESFS-----KOMRELETUKV---KOGGRPGATLEHQ 681
 Qy 862 QLEKISHQDSDKNRLLLELETRLRREVSLHEBEQKLEKRLQTELQLSQBPRESQLTALQA 921
 Db 682 EISKIRS-----ELEKKV---LFYEE---ELVREASHVLEVKVKEVHESES 724
 Qy 922 ARAALESQRLQAKTELESTHAEAEESI-QALTAHRDRIQKFKDALRNSCIVITDLESQLN 980

Db 725 HQALQKQVLMKQKLEKSKRHRSEMEBAIGAMKQKYER-----BRA 767
QY 981 QLTEDNAELNNQNFYLSKOLDASGAN---DEIVOLRSE---VDHLRREITEREMOLTS 1033
Db 768 MLFDENKKLTAENKELCSFVDKXLTQNRQLEDELDQSLASKESVAHWEAIAELIQWSD 827
QY 1034 QKQTMELKTTCTMLBEQWMDLEALNDELLEKERQWEAWRSVLGDEKSFQFCRVRBELQM 1093
Db 828 EKDARGYLQALASKMTE---ELETLRSSLSGRTLDPLWK-VRRSQKLDMSARL-ELQSA 882
QY 1094 LDTEKQSRARADQRTIESQVVELAVKEHKAETILAQALKEQKLAESLSKLANDLEK 1153
Db 883 LEAE-----IRAKQLVHEELRKVDTSLAFESKLKSEAK----- 917
QY 1154 HAMLENNARSLOOKJETERELKORLEEQAKLQQQMDLQKNHI FRLTQLOGLADRADLL 1213
Db 918 -----NRELLEENQSLKRMEEK---FRADTGL----- 942
QY 1214 KTERSLEYQLENIQVLYSHEKVKWEGTISQOTKLDFLOAKMDQPAKKKGLFSERKED 1273
Db 943 -----KLPDF-----QD 949
QY 1274 PALPTQVPLQVYNELKIALEKEKARCAELBEALQKTRIELRSAREBAHRKATDHPHPSTP 1333
Db 950 PIPE-----YFNTAPLAHDLTERTSSAQETQASKLDL-----SPSVVA 990
QY 1334 ATARQOIAMSAIVRSPHOPSAMSLAPPSSRRKSSSTPEBSRRLKERMHNNIHRNV 1393
Db 991 TSTEQO---EDARS-QORPSTVPL---ENTQALAMAGEKPKA-----HQRFS 1031
QY 1394 GLNMRATKCAVCLD-TVHFGROASKLECOVWCHPKCSTCLPATCGLPAEYATHTEAPC 1452
Db 1032 KGFPSPTQSHCTSLMGLIRGYACEVCAFCHVSKDSDAPQVPIPE----- 1081
QY 1453 RDKMNSPLQTKPE-----SSSLHLEGMKVPKRNKRQGGWDRKYVILEGSKVLIYD 1505
Db 1082 -----QSKRPLGVDVQRGIGTAYGYKVKVPKPT-GVKKGWGRAYAVVCDCKLFLYD 1131
QY 1506 -NEAREAGORPVEEFELCLPCDGVSHGAVGASELANATAKAKA---EADAKLLG----- 1556
Db 1132 LPEGKSTQGVFASQVLDLRDDEFV-SSVLASDVIHATRRDIPICFRVTASLGSPSKT 1190
QY 1557 NSLLKL---EGDDR-----LDMN-----CTLFPFSDQV----- 1580
Db 1191 SLLILITENENKRWVGILEGLQAILHKNRLRSQVHVHVAQYDSSFLIKTVLAAIV 1250
QY 1581 ----VLVGTESGLYALNVLKNLTHVPGIGAVFQIYIIXDLEKLMIAGERALCLVDVK 1636
Db 1251 DGDRIAVGLEEGLYVIELTRDVIIRAADCKYVQIELAPKEKLIILLCGRNH----- 1302
QY 1637 KVKQSLAQSHLPAQPDISPNIPEA-----VKGCHLFGAGKTENGLCICAAMPKSVV 1687
Db 1303 -----HVHLYPWTSPDGAENFIDKLPETKGCQOLIATGTRKSSSTCLFVAVKRL 1353
QY 1688 ILRYNNENLSKYCIRKEIETSEPCS-----CIHF-TNYSILIGNKTFYEDMKQY 1735
Db 1354 VLYCIETQRTKPHRFNETVAPGHVQVWAMFKDRLCVGVPSGFSLLSIQGDGQPLDL--- 1410
QY 1736 TLEEFILDKNDHSLAPAVFAASNSFPVSTVQVNSAGQREYVLLCHHEGVFVDSYGRSR 1795
Db 1411 -----VNPADPGLA---FLSQQSFALCAVELKS---EYLLCFSHMGLVYDPOGRSR 1458
QY 1796 TDDLKWSRLPLAFAYREPFLYFTHFNSLEVIPIQARSAGTPARAYLDIPNPRVYLGPAIS 1855
Db 1459 TQELAWPAAPVACSSSHVTVYSEYGVDFVDVTRME-----WVQTIGLRIRPLNS 1510
QY 1856 SCAIVLIASQDKLRVICCKGNLVKESGTEHHRGFTSRSS 1896
Db 1511 DGSNLLGC--EPPLIIVFKN---KFSGTVLN-VPTDSDNS 1545

RESULT 14

OS4875
ID 054875 PRELIMINARY; PRT; 1702 AA.
AC 054875;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myotonic dystrophy kinase-related Cdc42-binding kinase MRCK-beta.
GN MRCK-BETA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98078670; PubMed=9418861;
RA Leung T., Chen X.Q., Tan L., Manser E., Lim L.;
RT "Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a
RT Cdc42 effector in promoting cytoskeletal reorganization.";
RL Mol. Cell. Biol. 18:1130-140(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; AF021936; AAC02942.1; -.
DR PIR; T14050; T14050.
DR HSP; P28867; 1PTO.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR002219; DAG-PE-bind.
DR InterPro; IPR000095; FAKBox/RhoBindg.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG-PE-bind; 1.
DR Pfam; PF00786; PBD; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR ProDom; PD0000001; Foc_Kinase; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00285; PBD; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00108; CRIB; 1.
DR PROSITE; PS00479; DAG-PE BIND DOM_1; 1.
DR PROSITE; PS00081; DAG-PE BIND DOM_2; 1.
DR PROSITE; PS00003; PH DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1702 AA; 194019 MW; C7AF80707563D319 CRC64;

Query Match 13.1%; Score 1372; DB 11; Length 1702;
Best Local Similarity 24.7%; Pred. No. 1.5e-49;
Matches 476; Conservative 298; Mismatches 609; Indels 546; Gaps 65;
Qy 44 SPLSRGILDALFVFECSQPALMKIKHVSFVRKYSDTIALQLQPSAKDFVRSV 103
Db 23 SLSLVTLLDVLVCLYTECSHSLRRDKYVAEFLWAKPFTQLVQXQMLHREDFEIKVI 82
Qy 104 GCGHFAEVQVREKATGDIYAMKVMKKALLAQEQVFFEEERNILSRSTSPWIPQVYA 163

83 GRGAFGEVAVVMKNTERTIYAMKILNKWEMLKRAETACPREERDVLVNGDCQWITALHYA 142
164 FQDKNHLIMEEYQPGDILLNRRVEDOLDENLLOFYLAELILAVSHVHLGVRHDIK 223
143 FQDENLYLVMDYVVGDDLLTLLSKEDKLPEDMARFYIGEWLVAIDSLHQJHYVRHDIK 202
224 PENILVDRTHIKLVDPGSAKONSNKVMNAKLPICTPDYMPEVLTVNMGDKGTGID 283
203 PDNVLVDVNGHIRLADFGSLKWDGTOSSVAVGTPDYISPEILQAME-DGMGKYGPE 261
284 CDWWSGVGTAYEMYIGRSFPAECTSARTNNIMNFORFLKPPDD-PKVSDDFLDIQSLL 342
262 CDWWSLGVCMYEMLYGETPFYAESLVETYGKIMNHEERFQPSHVTDVSEBAKDILQIRLI 321
343 CGOKERLKPEGL---CHPFFSKIDNNIRNPPVPVPTLKSDDDTSNFDPE---KNSM 396
322 CSERRLQNGIEDFKHAFEGELAHENIRNLEAPIPDVSSPDSNFDVDDVLRNIE 381
397 VSSSPQOLSP---SGPSGELPVGFSYKALGILGRSSVSGLDSPAKTSMEKKLLI 453
382 I-----LPPGSHGTGSLHLPPIGFTT-----TESCFS--DRGSLKSMIQSNTLT 425
454 KSELODSQDKCHKMQEMTRLHRRYSEVEAVLSQKEVELKASQTORSLEQDLATYITE 513
426 KQDVQORDLENSIQIEAYERRR----- 449
514 CSSLKRSLQARMEVS---QEDDKALQLLHDIRE---QSRKLOEIKQEYQAOVEMRLM 567
450 -----LEQKLESLKQLQESTQVQSLHSGSTRALGNSNRDKETK--RLNEELERMKSK 500
568 M-----NOL-----EDLVARRSDIYSELRESLAAEFKPKATECOHLLKAKQCKPEV 621
501 MADSNRLERQLEDTVLRQ----- 519
622 GEYAKLEKINAEQOLKIQELQEKLEKAAKERAERLEKLQNRDSSGSEGRKKLVAERRR 681
520 -----EHEDSTQRLKGLKQYELAQEKEE-----LHKQLVEASER- 555
682 HSLKVKRLNWERENRDKDIQTKSQOIQOMADKILLEBKREAGVSAQHLEVLHK 741
556 -----LKSQTKELKDAHQKALQEFSS-----ELNERMAELRSQKQVSRQLR 599
742 QKQHYEETIKVLDNOLKDLADKETLENMORHBEAEHEKGIILSEOKAMINAMDSKIR 801
600 DKSEEVAMQKIDS-MRQDIRKSE-----KSKKELEAR----- 632
802 SLBQRIVELSEANKLAANGSLFTQNNKAQEMISELROCKPYLETQAGKLEAQNKRLEE 861
633 -LEDVAEASERKRLREHSESFS-----KQWERELTKV---KQGRGPGCATLEHQ 681
862 QLEKISHQDHSKRNRLLETRLREYSLSEHEQKLEK-----QLTELQSLQERESQLT 917
682 EISKIRS-----ELEKKV---LFYEELVRRERSHVLEKVNKKEVHESHSOL 727
918 ALQARAALLESQRLQARTELETTAEABEIEIQAHTARDEIQKFDALRNSCTVITDLEE 977
728 ALQKEVLMLDKLEKSK--RERHSEMEBAIGAM--KDKYER----- 764
978 QLNQLTEDNALNNQNYLSKQDLDEAGAN-----DRIVOLRSE---VDHLRREITEREMQ 1030
765 ERAMLFDENKLLTAENEKLCFSVDKLTQNRQLEDBLQDLAKGSEVAHWEAQIAETIOW 824
1031 LTSQKQTEALTKTCTMLEQVMDLEALNDELLEKEROEAWRSLVLDGSKQFECRVEL 1090
825 VSEKDGARGYLQALASQWTE---ELTLRSSLSGSRITLPLWK-VRRSQKLDMSARL-EL 879
1091 QRMLDTEKQSRARADORITESTQVVELAVKHEKAEITLALQOALKQKLEKSLDKLNDL 1150
880 QSALEAB-----IRAKQVHEELRKVKDTSLAFESKLKSEAK----- 917
1151 EKXHAMLENNARSLQOKLETERELKORLLEEQAKLOQVMDLQXNHIFRUTQGLQALDRA 1210
918 -----NRELLEWQSLKKEVEK-----FRADTGL----- 942

1211 DLLKTERSDELEYOLENTQVLYSHEKVKMEGTISQOTKLIDFLQAKMDQPAKKKKGLFSRR 1270
943 -----KLPDF----- 947
1271 KEDPALPTQVLPQVNEUKLALKEKAKCALEBALQTRIELRSAREEAAHRAKATDHPHP 1330
948 -QDPIFE-----YFNAPLAHDLTFRITSSASDQETQASKDL- 997
1331 STPATARQIAMSALVRSPEHQPSAMSLAPPSSRRKESSTPEEFSSRLKERMHNIPHR 1390
988 SVATSTEQ---EDAARS-QORSTVPL---PNTQALAMAGPKPKA-----HQ 1028
1391 FNVGLNVRATKCAVCLD-TVHFGQAKCLCEQVMCHPKCSTCLPATCGLPAEYATHFTE 1449
1029 FSIKSFSPPTQCSHCTSLMVGLIRQGVACEVCAFSCHVSCDSAPQVCPIPE----- 1081
1450 AFCEDKNSPGLQTKEP-----SSSLHLEGWVKVPRNNKRGQGWDRKYIVLEGSKVL 1502
1082 -----QSKRPLGVVDVORGIGTAYKGVKVPKPT-GVYKGMORAYAVVCDCKLF 1128
1503 IYD-NEAREAGQRFVEEFELCPDGVSIHGAVGASELANATAKAKA---EADAKLLG-- 1556
1129 LYDLPECKSTQPGVIASQVLDLRDDEFAV-SSVLASDVIHATRRDIPCIPTVATSLIGSP 1187
1557 ---NSILKL---EGDDR-----LDMN-----CTLPSDQV-- 1580
1188 SKTSSLLITENENKRWGILLEGLOAILHKRLRSQVHVHVAQEAVDSSPLIKTVLAA 1247
1581 -----VLVTGEBGLYALNVLNLSLTHVPGIGAVFIYIKLEKILMTAGEERALCLV 1633
1248 AIVDGDRIANGLEBGLVIELTRDVIIVRAADCKVYQIELAPKEXKILLCGRNH----- 1302
1634 DVKKVKSQASQSHLPAOPDISPNIFE-----VKGCHLFGAGKIENGLCICAAMPS 1684
1303 -----HVHLYPWTSTFDGAESNFDIKLPETKGCQOLIATGTLRKSSTCLFVAV 1350
1685 KWILRYNENLSKYCIKKEIETSEPCS-----CIHF-TNYSILITGNTKFEIDM 1732
1351 KRLVLCVEIQETKFFHKEFNEIAPGVHVNAMFKORLCVGYSPGFSLLSIQGGQPLDL 1410
1733 KQYTLBFLDKNDHSLAPAVFAASSNFPVSIYQVNSAGOREBYLLCFHFGFVDSYGR 1792
1411 -----VNPADPSLA---FLSQSPDALCAVELKS---EYLLCFSHMGLYVDPQGR 1455
1793 RSRTDDLKWSRLPLAFAYREFYLVTHFNSLVEIEIQARSSAGTPPARAYLDDIPNRYLGP 1852
1456 RSRTQELMNPAPVACSCSSSHVTYVEYGVDFVKTME-----WVQTIGLRRIRP 1507
1853 AISSGAIYL 1861
1508 LNSDGSINL 1516
RESULT 15
Q7TT50 PRELIMINARY; PRT; 1713 AA.
AC Q7TT50;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CDC42-binding protein kinase beta.
GN CDC42BPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.Q., Wu S.L., Cheng Z.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY277589; AAP34402.1; -.
KW Kinase.

SQ SEQUENCE 1713 AA; 194779 MW; 0D63976BAC61DDE1 CRC64;

Query Match 13.0%; Score 1363; DB 11; Length 1713;
Best Local Similarity 24.5%; Pred. No. 3,5e-49;
Matches 506; Conservative 314; Mismatches 681; Indels 562; Gaps 72;

QY 44 SPLREGILDALFVLFEECSQPALMKIKHVSFVKYSDTIAELQELQSAKDFEVRSLV 103
DB 23 SALSVELLDVLVCLVTECSHSAIRDKXVAFLEWAKPFTQLVXDMQLHREDFEI IKVI 82

QY 104 GGHFAFVQVREKATGDIYAMKWKKALLAQEQVSFFEEERNILSRSTSWIQLQVLA 163
DB 83 GGAFGEVAVVQVNRTERIYAMKLNKWEMLKRAETACFREERDVLVNGDCOMITAHYA 142

QY 164 FODKHLVMEBYQPGDGLLSLNRVEDOLDENLQFYLAELILAVSHVMGYVHRDIK 223
DB 143 FODENVYLXVMDYVYGGDLTLLSREDFKLPEDMARFYIGEMVLAIDSIHQHLYVHRDIK 202

QY 224 PENILVDRTGHILKVDFGAAAKNSNMKNYAKLPITGPDYMAPEVLTVNMGKGKTYGLD 283
DB 203 PDNVLLDVNGHILRADFGSLKXNDGTVQSSVAVGTPTYISPEILOAME-DGMGKYGE 261

QY 284 CDWSVGVIAYEMIXGRSPFAGTSARTENNIMNFORELKPDD-PKVSDDFLDIQSL 342
DB 262 CDWSLGVCMYEMLYGETFPYAESLVETGYCKLNHEERFQFFSHVTDVSEAKDILQRLI 321

QY 343 CQOKERLKEGEL---CCHPFFSKIDMNNIRNPPFPVPTLKSDDDTSNFDEPE---KNSW 396
DB 322 CSRERRLGQNGIEDFKKHAFFEGNLWENIRNLEAPVPDVSPSPSTNFDVDDMLRNE 381

QY 397 VSSSPCOLSP---SGFSBELPVGFSYSKALGILGRSESVUGLDSAKTSSMEKKILLI 453
DB 382 I-----LPGSHTFSGHLHLPFIQFTFT-----TESCSF---DRGSLKSMTOSNTLT 425

QY 454 KSKELQDSQCKKMEQEMTRHRRVSEVEAVLSQKEVELKASQTSRSLLEODLATYITE 513
DB 426 KDQVQDRLNSLQIEAYERRIR----- 449

QY 514 CSLSKRSLEQARMEVS---QEDDKALQLLHDIRE---QSRKLOEIKQEOYQAQVEMELM 567
DB 450 -----LQOEKLE-SRKLQESTQTQVSLHGSTRALGNSNRDKETK--RLNEELERMKSK 500

QY 568 M---NOLE---BDLVARSRLDYSELRESLAAAEFPKFKATECOHKLKAKDQKQPEV 621
DB 501 MADSNRLERQLDVTILRO-----EHEDSTHRL-----KGLEKQYRLAR----- 539

QY 622 GEYAKLEKINABOQLKIOBLQBLEKAAKERAERELEKLNQNRDSESGIRKCLVABEERR 681
DB 540 -----QEK-----EELHKQLVEASER- 555

QY 682 HSLNKVKLETMERRENFLKDDIQTKSQOIQOMADKILELEEKHREAQVSAQHLEVLK 741
DB 556 -----LK9QTKELKDAHQORRALQEFSS-----ELNRMSELRLKQKVSRLR 599

QY 742 QKEQHVE---BKIVLNDQIKDADKDTLENMQRHEEAEHKGKILSEQKAMINAMDS 798
DB 600 DREESMEVAMQKIDSMRQDLRSEKSRKELEARLEDAEAAEASKERKLREHSESPCKQWER 659

QY 799 KIRSLERQIVELSEANKLAANSSLTQRMNKAQEBMISLQOKQFYLETQACKLEAQNKR 858
DB 660 ELEAL-----KVQKGRGPGGAASEHQEITSKIR-----SELEKVLIFYEELVRRBA----- 706

QY 859 LEEQLEKISHODSKNRLLLETRLRVLSLBEHEQKLELKRQL-----TELQLSQERES 914
DB 707 -----SH-----VLEVNKVEKH--DSHQLALQKEVLMLKDKLEKSKRERHS 749

QY 915 QUTALOARAAALBESLQRAKTELEETAPAEIEIQALTAHRDEIQKFDALRNSCTVITD 974
DB 750 E-----MEEAIGTVKDYERERAMLFDENKKLTAENEKL-----CSFVDK 789

QY 975 LBEQLNQLTEDNAELNNQNFYLSKQLDASGANDEIVQLRSEVDHLRREITERENQLTSQ 1034
DB 790 LTAQNRQLEDELODLASK-----KESVAHWEAQIAEIIQWVSDE 828

QY 1035 KQTWEALKTTCTMLLEEQVMDLEALNDELLEKQERQWEARWSVLGDEKQSFQEFQRELOML 1094
DB 829 KDARGYQALASKVTE---ELETURSSLSGSRTLDPLWK--VRRSQKLDMAAIL--ELQSAL 883

QY 1095 DTEQSRARADORITESQVQVVELAKHEKABILALQALKEQKLKABESLSDKLNLEFKKH 1154
DB 884 EAE-----IRAKQLVQELRKVRKVDSSLAFAESKLKSEAK----- 917

QY 1155 AMLEKNARSLOQKLETERELKQRLLEBQAKLOQOMDLOKNHIFRLTOGLOEALDRALLK 1214
DB 918 -----NRELLEEMQSLRKMEBK-----FRADTGLK-----L 944

QY 1215 TERSDLEYQLENIQVLYSHEKVKMEGTISOQTKLIDFLQAKMDQPAKKKXGLFSRRKEDP 1274
DB 945 PDFODSIFEYENTAPLAHDLTFRSSASDQETQ-----ASKMDLSPSVSVATSTEQEDM 999

QY 1275 ALPTQVLOYNELKLALEKEKARCAELEBALQKTRIBLSAREBAHUKATDHPHPTPA 1334
DB 1000 ARPOQRP-----SPVFLPSTQA 1016

QY 1335 TARQOIAMSAIVRSPEHQPSAMSLLAPPSRRRKESSTPEEFSSRLKERMHHNTPHRNVG 1394
DB 1017 LA-----MAGPKPK-----AHQPSIK 1032

QY 1395 LNMKATKAVCLD-TVHFGRAQSKLEQVNVCHPKCSTCLPATCGLPAEYATHFTFAFCR 1453
DB 1033 SFPSPQCSHCTSLMWGLIRQGYACEVCAFSCHVSCXDSAPQVCPPIPE----- 1081

QY 1454 DKNSPGLQKEP-----SSSLHLEGMKVRNNKRGQGWDRKYIVLEGSKVLIYD- 1505
DB 1082 -----QSRPLGVQVQRGIGTAYKGYVKPKPT-GVKKGWQRAYAVVCDCKFLFYDL 1132

QY 1506 NEAREAQRPVEEPELCLPDGDVSIHGAVGASELANATAKAKA---BADAKLLG---N 1557
DB 1133 PEGKSTQGVVASQVLDLRDEEFAV--SSVLASDVIAHTRDIPCIPTVATSLGSPSKTS 1191

QY 1558 SLLKLEGGDRIDMNCNLTLPFSDQVVLVGTETEGVAL---NVLKNSLTHV-----PGI 1605
DB 1192 SLLTLTENE-----NEKRWGILEGLQAILHKNRLKQVHVHVAQEAVIDSSLEPI 1241

QY 1606 GAVFOIYIIKDLKMLMIAEERALCLV-----DVKKV-----KOSLA--- 1643
DB 1242 XAVLAAIV-DGRIAV--GLEEGLYVIELTRDVIIVRAADCKKVQYQIELAFKELIALLC 1298

QY 1644 --QSHLPAQPDIDSNITFEA-----VKGCHLFGAGKIENGLICICAAFPSKVILRYN 1692
DB 1299 GRNHVHLYPWSSPDGAASNFIDKLPETKGCQLIATGLRKSSSTCLFVAVKRLILCYE 1358

QY 1693 ENLSKYCIRKEIETSEPCS-----CIHF-TNYSILIGTNKEYEIDMKQYTLLEF 1740
DB 1359 IQRTKPHRKFSELVAPGHVQWNAVFKDRLCVGPSPGFSLSLSTQGGDPPDL----- 1410

QY 1741 LDKNDHSLAPAVFAAASSNSFPVSIQVNSAQOREYLLCFHEFGVFVDSYGRSRTDDLK 1800
DB 1411 VNPTDPSLA---FLSQSQSPDALCAVELKS---EYLLCFSHMGLYVDPQGRSRMQEML 1463

QY 1801 WSRPLAFAYREPYLFVTHFNSLEVEIQARSAGCTPARAYLDIPNPRYLGPAISSAIY 1860
DB 1464 WPAAPVACSCSPHTVTYSEYGVDPDVTME-----WVQTIGLRIRPRLNSDGLN 1515

QY 1861 LASSYQDKLAVICCKGNLVKESGTEHHRGPSTGRSPNKRGPPTYNEHITKRVASSPAPP 1920
DB 1516 LLGC--EPPELLIYFKN---KFSGTILN-VPDTSNKKQMLRTSKERFVKV----- 1562

QY 1921 EGSHPREPSTPHRYEGRTELR-----RDKSPORPLE-----REKSPGRMLSTRE--RSP 1970
DB 1563 -----PEE-----ERLQORREMLRDPRLRSKVISNPTNFNVHVMGPGDGMQVLMDLPLSA 1613

QY 1971 GRUFEDSSRCRLPAGAVRTPLSQ 1993
DB 1614 APTVQBEKQGTTPAGLPRQPPSR 1636

Search completed: July 3, 2004, 10:17:33
Job time : 113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 20:02:52 ; Search time 14333 Seconds
(without alignments)
13696.642 Million cell updates/sec

Title: US-10-017-216-1
Perfect score: 6574
Sequence: 1 agagcgcagtgaggagat.....atcgagaatgaggtttaga 6574

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_esthum.*
- 4: em_esthum.*
- 5: em_esthum.*
- 6: em_esthum.*
- 7: em_esthum.*
- 8: em_esthum.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_est3.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	915.2	13.9	1011	13 BQ070955	BQ070955 AGENCOURT
2	815.6	12.4	956	13 BQ071141	BQ071141 AGENCOURT
3	810.2	12.3	830	13 BUI181633	BUI181633 AGENCOURT
4	750.6	11.4	849	12 B1253509	B1253509 602973370

5	716.4	10.9	920	13	BX342268	BX342268
6	692	10.5	879	13	BQ228524	BQ228524 AGENCOURT
7	688.8	10.5	958	11	EC031156	EC031156 Mus muscu
c	629.2	9.6	652	10	BF905370	BF905370 IL3-MT026
9	606.6	9.2	881	12	BG976452	BG976452 602846269
10	589.8	9.0	746	14	CA749290	UI-M-FY0-
11	584.6	8.9	843	12	BG912161	602812833
12	573.6	8.7	647	14	CF744580	UI-M-GVO-
13	572.8	8.7	730	14	CF723360	CF723360 UI-M-GVO-
c	570	8.7	826	14	CD172337	AGENCOURT
c	570	8.7	894	13	BQ421272	AGENCOURT
c	570	8.7	2426	11	BC018818	Homo sapi
17	566.4	8.6	1085	12	BM904785	BM904785 AGENCOURT
18	557.8	8.5	1201	13	BX449104	BX449104
19	556.8	8.5	640	13	BQ807302	BQ807302 NISC_kk01
20	555.6	8.5	963	13	BX437603	BX437603
c	544.8	8.3	778	10	BE910475	BE910475 601503793
c	540	8.2	540	9	AA746301	AA746301 oa56C11.i
c	534	8.1	805	14	CF126292	CF126292 UI-HF-ET0
c	523	8.0	726	10	BE735426	BE735426 601304218
25	520.6	7.9	1044	9	AL532373	AL532373 AL532373
c	514	7.8	854	13	BUI195960	BUI195960 AGENCOURT
27	511.8	7.8	775	13	BUI105833	BUI105833 603005490
c	511.4	7.8	954	12	BG513123	BG513123 602811735
29	511	7.8	804	13	BUI236617	BUI236617 603411670
30	510.4	7.8	933	13	BUI232508	BUI232508 603408272
31	508.6	7.7	547	14	CF744826	CF744826 UI-M-GVO-
c	507	7.7	564	12	BM826567	BM826567 K-EST0098
33	506.6	7.7	965	9	AL580381	AL580381 AL580381
c	502	7.6	940	12	BG744767	BG744767 602722887
c	497.4	7.6	572	9	AA034099	AA034099 z106108.i
c	492	7.5	624	9	AA044766	AA044766 zk67D10.i
37	491.6	7.5	875	13	BX723780	BX723780 BX723780
38	490.8	7.5	875	13	BUI900029	BUI900029 AGENCOURT
c	488.8	7.4	1127	12	BM544893	BM544893 AGENCOURT
40	484.2	7.4	820	13	BQ941506	BQ941506 AGENCOURT
41	473	7.2	647	10	BE207065	BE207065 BE207065
42	472.6	7.2	657	13	BQ831488	BQ831488 LL61n2149
43	471.6	7.2	638	14	CD348416	CD348416 UI-M-FY0-
44	471.4	7.2	1001	9	AL579049	AL579049 AL579049
c	466.6	7.1	681	10	AW605350	AW605350 QV3-DT004

ALIGNMENTS

RESULT 1
BQ070955
LOCUS BQ070955 1011 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6855847 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5923441
S., mRNA sequence.
ACCESSION BQ070955
VERSION BQ070955.1 GI:19900001
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1011)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2091 row: p column: 02
High quality sequence stop: 634.

```

FEATURES
  source      Location/Qualifiers
              1..1011
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5923441"
                /tissue_type="neuroblastoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_47"
                /notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGACAGG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."

ORIGIN
Query Match      13.9%; Score 915.2; DB 13; Length 1011;
Best Local Similarity 97.9%; Pred. No. 2.1e-208;
Matches 946; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

QY 2671 CGGAGGTGAGTCTAGACGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2730
DB 1 CGGAGGTGAGTCTAGACGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 60

QY 2731 CTACAGCTCTCCCTGAGGAGCGCGAGTCAAGTTCAGAGCCCTGAGGCTCGAGGGG 2790
DB 61 CTACAGCTCTCCCTGAGGAGCGCGAGTCAAGTTCAGAGCCCTGAGGCTCGAGGGG 120

QY 2791 GCCTGGAGGCGAGCTTGGCGAGGCGAGAGACAGAGCTGGAGAGACACACAGAGCT 2850
DB 121 GCCTGGAGGCGAGCTTGGCGAGGCGAGAGACAGAGCTGGAGAGACACACAGAGCT 180

QY 2851 GAAGAGGAGATCAGGACCTCAGGCGACATAGAGATGAATCCAGCGCAAAATTGATGT 2910
DB 181 GAAGAGGAGATCAGGACCTCAGGCGACATAGAGATGAATCCAGCGCAAAATTGATGT 240

QY 2911 CTTGCTAACAGCTGTACTGTATATACAGACCTGGAGAGCAGCTTAACAGCTGACCGAG 2970
DB 241 CTTGCTAACAGCTGTACTGTATATACAGACCTGGAGAGCAGCTTAACAGCTGACCGAG 300

QY 2971 GACAACTGTAACCTCAACACCAAACTTCTACTTCTCCAAACAACTCGATGAGGCTTCT 3030
DB 301 GACAACTGTAACCTCAACACCAAACTTCTACTTCTCCAAACAACTCGATGAGGCTTCT 360

QY 3031 GCGGCCAAACGACGAGATTTGACAACTGCGAAGTGAAGTGAAGCATTCTCCCGCGGAGATC 3090
DB 361 GCGGCCAAACGACGAGATTTGACAACTGCGAAGTGAAGTGAAGCATTCTCCCGCGGAGATC 420

QY 3091 ACGGACGAGAGATGAGCTTACAGCCAGAACCAACAGTGAAGCTCTGAAGACCAAG 3150
DB 421 ACGGACGAGAGATGAGCTTACAGCCAGAACCAACAGTGAAGCTCTGAAGACCAAG 480

QY 3151 TGCACCATGCTGGAGGAACAGAGTTCATGGATTGGAGGCGCTTAAACGATGAGTGTAGAA 3210
DB 481 TGCACCATGCTGGAGGAACAGAGTTCATGGATTGGAGGCGCTTAAACGATGAGTGTAGAA 540

QY 3211 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGTCTGGTGTATCAGAAATCCAGTTTGA 3270
DB 541 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGTCTGGTGTATCAGAAATCCAGTTTGA 600

QY 3271 TGTGCGGTTTCGAGAGTTCGAGAGAAATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3330
DB 601 TGTGCGGTTTCGAGAGTTCGAGAGAAATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 660

QY 3331 GATCAGCGGATACCGAGTCTCCAGGTGCTGGAGCTGGAGTGGAGGAGCAGACAGGCT 3390
DB 661 GATCAGCGGATACCGAGTCTCCAGGTGCTGGAGTGGAGGAGCAGACAGGCT 720

QY 3391 GAGATTCTCGCTCTGACGAGCGGCTCTCAAAGAGCAGAGCTGAAGGCCGAGAGCCTCTCT 3450

```

```

DB 721 GAGATTCTCGCTCTGACGAGCGGCTCTCAAAGACAGAGCTGGAGCCGAGAGCCTCTCT 780
QY 3451 GACAAGCTCAATGACCTGGAGAGAGAGCATGTATGCTTGAATGAATGCC- GAAGCTT 3509
DB 781 GACAAGCTCAATGACCTGGAGAGAGAGCATGTATGCTTGAATGAATGCCCGAAGCTT 840
QY 3510 ACACGAGAAGCTGGAGACTCAAGAGAGCTCAACAGAGGCTTCTGAGAGAGCAAGCCAA 3569
DB 841 ACACGAGAAGCTGGAGACTGGAGAGCTCAACAGAGGCTCTGGANANCAANCCAA 900
QY 3570 ATTACACGACGATGGACCTGCAGAAAAATCACTTTT--CGTCTGACTCAAGGACTG 3627
DB 901 ATTACACGACGATGGATGGATGGCGAAATCACTATTTTCGCGTGGACTCAAGGACTG 960
QY 3628 CAAGAA 3633
DB 961 CCAGAA 966

RESULT 2
BO071141
LOCUS      BO071141
DEFINITION AGENCOURT_6853098 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927502
           5', mRNA sequence.
ACCESSION  BO071141
VERSION    BO071141.1 GI:19900187
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 956)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-x@mail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L1CM2102 row: i column: 07
           High quality sequence stop: 650.

FEATURES
  source      Location/Qualifiers
              1..956
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5927502"
                /tissue_type="neuroblastoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_47"
                /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGACAGG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."

ORIGIN
Query Match      12.4%; Score 815.6; DB 13; Length 956;
Best Local Similarity 97.1%; Pred. No. 1.9e-184;
Matches 874; Conservative 0; Mismatches 19; Indels 7; Gaps 4;

QY 2671 CGGAGGTGAGTCTAGAGCAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2730

```

Db 1 CGGGAGGTGCTAGACGACGAGGAGCAGAACTGGAGCTCAAGGCGAGCTCACAG 60
QY 2731 CTACAGCTCCCTGCAGGAGCGGAGTCACAGTTTACAGCCCTGCAGGCTGCAGGCG 2790
Db 61 CTACAGCTCCCTGCAGGAGCGGAGTCACAGTTTACAGCCCTGCAGGCTGCAGGCG 120
QY 2791 GGCCTGGAGAGCCAGCTTCGCCAGGCGGAGACAGAGCTGGAGAGACACACAGAACT 2850
Db 121 GGCCTGGAGAGCCAGCTTCGCCAGGCGGAGACAGAGCTGGAGAGACACACAGAACT 180
QY 2851 GAAGAGGAGATCCAGGCACTTCAGGCGACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910
Db 181 GAAGAGGAGATCCAGGCACTTCAGGCGACATAGAGATGAATCCAGCGCAAAATTTGATGCT 240
QY 2911 CTTGCTAACAGCTGTACTGTATTCACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2970
Db 241 CTTGCTAACAGCTGTACTGTATTCACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 300
QY 2971 GACAAAGCTGAACTCAACAAACAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030
Db 301 GACAAAGCTGAACTCAACAAACAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 360
QY 3031 GCGCGCAACGAGAGATTGTAACAATGGGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3090
Db 361 GCGCGCAACGAGAGATTGTAACAATGGGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 420
QY 3091 ACGGAACGAGAGATGAGCTTACCGAGCAGAGAGCAACGATGGAGGCTCTGAGACACAG 3150
Db 421 ACGGAACGAGAGATGAGCTTACCGAGCAGAGAGCAACGATGGAGGCTCTGAGACACAG 480
QY 3151 TGCACCATGCTGGAGGAACAGGTCTATGGATTGGAGGCCCTAAAGCATGAGCTGTGAA 3210
Db 481 TGCACCATGCTGGAGGAACAGGTCTATGGATTGGAGGCCCTAAAGCATGAGCTGTGAA 540
QY 3211 AAAGAGCGGAGCTGGAGGCCCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3270
Db 541 AAAGAGCGGAGCTGGAGGCCCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 600
QY 3271 TGTCTGGTTCGAGAGCTGCAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3330
Db 601 TGTCTGGTTCGAGAGCTGCAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 660
QY 3331 GATCAGCGGATCACCAGGCTCCGAGGTGTTGGAGCTGGAGTGAAGAGGACAAAGCT 3390
Db 661 GATCAGCGGATCACCAGGCTCCGAGGTGTTGGAGCTGGAGTGAAGAGGACAAAGCT 720
QY 3391 GAGATTCTCGCTTCGAGAGGCTCTCAAGAGCAGAGCTGAGGCGGAGAGCTCTCT 3450
Db 721 GAGATTCTCGCTTCGAGAGGCTCTCAAGAGCAGAGCTGAGGCGGAGAGCTCTCT 780
QY 3451 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCGCGAAGC 3507
Db 781 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCGCGAAGC 840
QY 3508 TTACAGCAGAA--GCTGGAGACTGAAC--GAGAGCTCAACAGAGGCTTCTGGAAGACA 3563
Db 841 TTACAGCAGAAAGCTGGGAGACTGAACCGAGAAGCTCAACAGAGGGTTCTTGGGAAGGA 900

RESULT 3
BU181633
LOCUS
DEFINITION AGENCOURT_7906225 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6140538
5', mRNA sequence.
ACCESSION BU181633
VERSION BU181633.1 GI:22695617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC <http://mgi.mgi.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13459 row: e column: 19
High quality sequence stop: 652.
FEATURES
source
1. 830
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6140538"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN
Query Match 12.3%; Score 810.2; DB 13; Length 830;
Best Local Similarity 99.4%; Pred. No. 3.4e-183;
Matches 823; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 2405 TGGATTCACAGATCAGATCCCTGGACAGAGGATTTGGAACTGTCTGAAAGCCAAATAAAC 2464
Db 1 TGGATTCACAGATCAGATCCCTGGACAGAGGATTTGGAACTGTCTGAAAGCCAAATAAAC 60
QY 2465 TTGCAGCAATAGCAGCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTT 2524
Db 61 TTGCAGCAATAGCAGCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTT 120
QY 2525 CTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGCTGGAGTTGGAGGCCCAGA 2584
Db 121 CTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGCTGGAGTTGGAGGCCCAGA 180
QY 2585 ACCGAAATCTGGAGGAGCAGCTGGAGAGATCAGCCACCAAGACACAGTGAAGAAATC 2644
Db 181 ACCGAAATCTGGAGGAGCAGCTGGAGAGATCAGCCACCAAGACACAGTGAAGAAATC 240
QY 2645 GGCTGCTGGAATCTGGAGACAAAGATTGCGGAGGTCTAGTCTAGAGACGAGGACGAAAC 2704
Db 241 GGCTGCTGGAATCTGGAGACAAAGATTGCGGAGGTCTAGTCTAGAGACGAGGACGAAAC 300
QY 2705 TGGAGCTCAAGGCCAGCTCAGAGCTACAGCTCTCCCTGCGAGGAGCGGAGTCAAGT 2764
Db 301 TGGAGCTCAAGGCCAGCTCAGAGCTACAGCTCTCCCTGCGAGGAGCGGAGTCAAGT 360
QY 2765 TGAAGCCCTTGCAGGCTGACCGGCGGCTGGAGAGCAGCTTCCGCCAGGCGAAGACAG 2824
Db 361 TGAAGCCCTTGCAGGCTGACCGGCGGCTGGAGAGCAGCTTCCGCCAGGCGAAGACAG 420
QY 2825 AGCTGGAGAGCAGCAGAGAGCTGAGAGAGATCCAGGACCTCAGGCACTCAGGCACTAGAG 2884
Db 421 AGCTGGAGAGCAGCAGAGAGCTGAGAGAGATCCAGGACCTCAGGCACTCAGGCACTAGAG 480
QY 2885 ATGAATTCAGGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCAGAGACTGG 2944
Db 481 ATGAATTCAGGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCAGAGACTGG 540
QY 2945 AGGAGAGCTAAACCGAGTGAACCGAGACAACTGAACTCAACCAACCAAACTTCTACT 3004
Db 541 AGGAGAGCTAAACCGAGTGAACCGAGACAACTGAACTCAACCAACCAAACTTCTACT 600
QY 3005 TGTCCAAACAACTCGATGAGGCTTCTGGGCGCCAAACGACGAGATTGTACAACTCGGAAGTG 3064

```

Db      601 TGTCCAAACACTCGATGAGGCTCTTGGCCCAACGACGAGATTGTACACTGCGAAGTG 660
QY      3065 AAGTGACCATCTCCCGCGGAGATCAGCAACGAGAGATGAGCTTACCACCGAAGC 3124
Db      661 AAGTGACCATCTCCCGCGGAGATCAGCAACGAGAGATGAGCTTACCACCGAAGC 720
QY      3125 AAGCATGAGGCTCTGAAGACCACTGACCATGCTGAGGAAACAGGTCATGGATT-TG 3183
Db      721 AAGCATGAGGCTCTGAAGACCACTGACCATGCTGAGGAAACAGGTCATGGATTGG 780
QY      3184 GAGGCCCTAAACGATGAGCTGTAGAAAAGAGCGGCACTGGAGGCC 3231
Db      781 GAGGCCCTAAGCATGAGCTGTAGAAAAGAGCGGCACTGGAGGCC 828

RESULT 4
BI253509 849 bp mRNA linear EST 17-JUL-2001
LOCUS 602973370F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112737 5',
DEFINITION mRNA sequence.
ACCESSION BI253509
VERSION 1
KEYWORDS GI:14805003
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 849)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgs.nhl.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11273 row: 1 column: 18
High quality sequence start: 2
High quality sequence stop: 772.
Location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5112737"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

FEATURES
source
Query Match 11.4%; Score 750.6; DB 12; Length 849;
Best Local Similarity 98.5%; Pred. No. 7.7e-169;
Matches 789; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY      4677 ACTGCTTGAACCTCCCTGCTGAACTGGAAGTGATGACCGCTAGACATGAATGAC 4736
Db      3 ACAGCTCGGAATCTCCCTGCTGAACTGGAAGTGATGACCGCTAGACATGAATGAC 62
QY      4737 GCTGCCCTTCAGTGACCACTGCTGTGTGGTGGCCACCGAGGAGGCTTACGCCCTGAA 4796
Db      63 GCTGCCCTTCAGTGACCACTGCTGTGTGGTGGCCACCGAGGAGGCTTACGCCCTGAA 122
QY      4797 TGCTTTGAAAACCTCCCTAACCATCTCCAGGAATGGAGGAGTCTTCCAAATTATAT 4856

```

```

Db      123 TGTCTTGAATAATCCCTCAACCATCTGCCAGGAATCGGAGCAGTCTTCCAAATTTATAT 182
QY      4857 TATCAAGAGACTGAGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGT 4916
Db      183 TATCAAGAGACTGAGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGT 242
QY      4917 GGACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACAT 4976
Db      243 GGACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACAT 302
QY      4977 CTACCCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCAGCGAAGATTGA 5036
Db      303 CTACCCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCAGCGAAGATTGA 362
QY      5037 GAACGGGCTCTGTCATCTGTGAGCCATGCCAGCAAGTCTCATTTCCGCTACACGA 5096
Db      363 GAACGGGCTCTGTCATCTGTGAGCCATGCCAGCAAGTCTCATTTCCGCTACACGA 422
QY      5097 AAACCTCAGCAATATCTGTCATCCGAAAGAGATAGACCTCAGACCTTGCAGCTGTAT 5156
Db      423 AAACCTCAGCAATATCTGTCATCCGAAAGAGATAGACCTCAGACCTTGCAGCTGTAT 482
QY      5157 CCATTTCACCAATTACAGTATCTCTTGGAAACCAATAAATTTCTAGAAATCGACATGAA 5216
Db      483 CCATTTCACCAATTACAGTATCTCTTGGAAACCAATAAATTTCTAGAAATCGACATGAA 542
QY      5217 GCAGTACACGCTCAGGAAATTCCTGGATAAAGATGACCATTCCTTGGCACCTGTGTGT 5276
Db      543 GCAGTACACGCTCAGGAAATTCCTGGATAAAGATGACCATTCCTTGGCACCTGTGTGT 602
QY      5277 TGCCGCTCTTCCAAACAGCTTCCTCTGTCCTCAATCTGTCAGGTGAACAGCGCAGGAGCG 5336
Db      603 TGCCGCTCTTCCAAACAGCTTCCTCTGTCCTCAATCTGTCAGGTGAACAGCGCAGGAGCG 662
QY      5337 AGAGGAGTACTTGTCTGTGTTTCCACGAATTTGAGTGTTCGTGGATTCTTACGGAAGACG 5396
Db      663 AGAGGAGTACTTGTCTGTGTTTCCACGAATTTGAGTGTTCGTGGATTCTTACGGAAGACG 722
QY      5397 T-AGCCGACAGAGATCTCAAGTGGAGTCTTACCTTTGGCTTGGCTTACGAGAAC 5455
Db      723 TAAGCCGACAGAGATCTCAAGTGGAGTCTTACCTTTGGCTTGGCTTACGAGAAC 780
QY      5456 CCTATCTGTTTGTGACCCACT 5476
Db      781 CCTATCTGTTTGGAGCCACT 801

RESULT 5
BI253509 920 bp mRNA linear EST 02-MAY-2003
LOCUS BX342268
DEFINITION BX342268 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK011VB20 5-PRIME, mRNA sequence.
ACCESSION BX342268
VERSION BX342268.1 GI:30334095
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 920)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6533.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK011DA10QPL&cluster=6533.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

```

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK011DA10QP1.

FEATURES
SOURCE

```

1. 920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK011YB20"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	10.9%;	Score 716.4;	DB 13;	Length 920;
Best Local Similarity	95.9%;	Pred. No. 1.4e-160;		
Matches 746;	Conservative 0;	Mismatches 31;	Indels 1;	Gaps 1;
QY	5283	CTCTTCCACAGCTTCCCTGTCTCAATCGTCGACGGTGAACAGCGCAGGGCAGCGAGAGGA	5342	
DB	66	CTCTTCCACAGCTTCCCTGTCTCAATCGTCGACGGTGAACAGCGCAGGGCAGCGAGAGGA	125	
QY	5343	GTACTTGCCTGTGTTTCCACGAATTTGAGAGTTCGTGGAATTCCTACGAAAGAGCTAGCCG	5402	
DB	126	GTACTTGCCTGTGTTTCCACGAATTTGAGAGTTCGTGGAATTCCTACGAAAGAGCTAGCCG	185	
QY	5403	CACAGACGATCTCAAGTGGAGTCGTACCTTTGGCCCTTTGCCCTACAGAGAACCCCTATCT	5462	
DB	186	CACAGACGATCTCAAGTGGAGTCGTACCTTTGGCCCTTTGCCCTACAGAGAACCCCTATCT	245	
QY	5463	GTTTGTGACCCACTTCAAACTCACTCGAAGTAATTGAGATCCAGGACGCTCCCTCAGCAGG	5522	
DB	246	GTTTGTGACCCACTTCAAACTCACTCGAAGTAATTGAGATCCAGGACGCTCCCTCAGCAGG	305	
QY	5523	GACCCCTGCCGAGCGTA CTTGACATCCCGAACC CGCGCTA CTTGGGCCCTGCCAATTC	5582	
DB	306	GACCCCTGCCGAGCGTA CTTGACATCCCGAACC CGCGCTA CTTGGGCCCTGCCAATTC	365	
QY	5583	CTCAGAGCGATTTACTTTGCGTCCTCATACACAGGATAAATTAAGGGTCATTTCGTGCAA	5642	
DB	366	CTCAGAGCGATTTACTTTGCGTCCTCATACACAGGATAAATTAAGGGTCATTTCGTGCAA	425	
QY	5643	GGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCAACCGGGGCCGCTCACTCCCGCAG	5702	
DB	426	GGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCAACCGGGGCCGCTCACTCCCGCAG	485	
QY	5703	CAGCCCCAACAGCAGGGCCCCACCCAGTACACAGGACATACACAGCGCGTGGCCTC	5762	
DB	486	CAGCCCCAACAGCAGGGCCCCACCCAGTACACAGGACATACCAAGCGCGTGGCCTC	545	
QY	5763	CAGCCAGCGCCGCCCGAAGGCCCGCAGCCACCCCGCAGAGCCAAAGCACACCCACCGCTA	5822	
DB	546	CAGCCAGCGCCGCCCGAAGGCCCGCAGCCACCCCGCAGAGCCAAAGCACACCCACCGCTA	604	
QY	5823	CCGCGAGGGCGGACCCGAGTCTGCGAGGGACAAGTCTCTGGCCGCCCTGAGAGGAGA	5882	
DB	605	CCGCGAGGGCGGACCCGAGTCTGCGAGGGACAAGTCTCTGGCCGCCCTGAGAGGAGA	664	
QY	5883	GAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGGTCTCCCGGAGGCTGTTGA	5942	
DB	665	GAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGGTCTCCCGGAGGCTGTTGA	724	
QY	5943	AGACAGCAGCAGGGGCCGCTGCTCGGGAGCGCTGAGGACCCCGCTGTCCCGAGGTGAA	6002	
DB	725	AGACAGCAGCAGGGGCCGCTGCTCGGGAGCGCTGAGGACCCCGCTGTCCCGAGGTGAA	784	
QY	6003	CAGGGAAGGCGAGTGCCTCTCAAGTTTTCACGGTTTAACTGTCACCTATTAT	6060	
DB	785	CAGGCTCTGGGACGAGTCTTCAGTATATAATCTCAGCCAGAAAAACCAATCTCCTCATCT	842	

RESULT 6
BQ228524
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Employment Status</i>	11. <i>Employment Status</i>
12. <i>Family Size</i>	12. <i>Family Size</i>
13. <i>Home Ownership</i>	13. <i>Home Ownership</i>
14. <i>Vehicle Ownership</i>	14. <i>Vehicle Ownership</i>
15. <i>Insurance Status</i>	15. <i>Insurance Status</i>
16. <i>Subscription Services</i>	16. <i>Subscription Services</i>
17. <i>Charitable Donations</i>	17. <i>Charitable Donations</i>
18. <i>Volunteer Work</i>	18. <i>Volunteer Work</i>
19. <i>Real Estate Transactions</i>	19. <i>Real Estate Transactions</i>
20. <i>Financial Investments</i>	20. <i>Financial Investments</i>
21. <i>Medical History</i>	21. <i>Medical History</i>
22. <i>Travel Plans</i>	22. <i>Travel Plans</i>
23. <i>Employment History</i>	23. <i>Employment History</i>
24. <i>Family Income</i>	24. <i>Family Income</i>
25. <i>Home Address</i>	25. <i>Home Address</i>
26. <i>Vehicle Registration</i>	26. <i>Vehicle Registration</i>
27. <i>Insurance Claims</i>	27. <i>Insurance Claims</i>
28. <i>Subscription Cancellations</i>	28. <i>Subscription Cancellations</i>
29. <i>Charitable Receipts</i>	29. <i>Charitable Receipts</i>
30. <i>Volunteer Hours</i>	30. <i>Volunteer Hours</i>
31. <i>Real Estate Listings</i>	31. <i>Real Estate Listings</i>
32. <i>Investment Returns</i>	32. <i>Investment Returns</i>
33. <i>Medical Records</i>	33. <i>Medical Records</i>
34. <i>Travel Itineraries</i>	34. <i>Travel Itineraries</i>
35. <i>Employment Records</i>	35. <i>Employment Records</i>
36. <i>Family Expenses</i>	36. <i>Family Expenses</i>
37. <i>Home Mortgage</i>	37. <i>Home Mortgage</i>
38. <i>Vehicle Leasing</i>	38. <i>Vehicle Leasing</i>
39. <i>Insurance Policies</i>	39. <i>Insurance Policies</i>
40. <i>Subscription Renewals</i>	40. <i>Subscription Renewals</i>
41. <i>Charitable Tax Deductions</i>	41. <i>Charitable Tax Deductions</i>
42. <i>Volunteer Certificates</i>	42. <i>Volunteer Certificates</i>
43. <i>Real Estate Appraisals</i>	43. <i>Real Estate Appraisals</i>
44. <i>Investment Statements</i>	44. <i>Investment Statements</i>
45. <i>Medical Prescriptions</i>	45. <i>Medical Prescriptions</i>
46. <i>Travel Vouchers</i>	46. <i>Travel Vouchers</i>
47. <i>Employment Contracts</i>	47. <i>Employment Contracts</i>
48. <i>Family Budgets</i>	48. <i>Family Budgets</i>
49. <i>Home Renovation Plans</i>	49. <i>Home Renovation Plans</i>
50. <i>Vehicle Maintenance Records</i>	50. <i>Vehicle Maintenance Records</i>
51. <i>Insurance Claims History</i>	51. <i>Insurance Claims History</i>
52. <i>Subscription Usage Data</i>	52. <i>Subscription Usage Data</i>
53. <i>Charitable Impact Reports</i>	53. <i>Charitable Impact Reports</i>
54. <i>Volunteer Training Records</i>	54. <i>Volunteer Training Records</i>
55. <i>Real Estate Market Analysis</i>	55. <i>Real Estate Market Analysis</i>
56. <i>Investment Performance Metrics</i>	56. <i>Investment Performance Metrics</i>
57. <i>Medical Test Results</i>	57. <i>Medical Test Results</i>
58. <i>Travel Itinerary Changes</i>	58. <i>Travel Itinerary Changes</i>
59. <i>Employment Termination Notices</i>	59. <i>Employment Termination Notices</i>
60. <i>Family Financial Statements</i>	60. <i>Family Financial Statements</i>
61. <i>Home Energy Audit Reports</i>	61. <i>Home Energy Audit Reports</i>
62. <i>Vehicle Recall Notifications</i>	62. <i>Vehicle Recall Notifications</i>
63. <i>Insurance Policy Updates</i>	63. <i>Insurance Policy Updates</i>
64. <i>Subscription Service Reviews</i>	64. <i>Subscription Service Reviews</i>
65. <i>Charitable Grant Applications</i>	65. <i>Charitable Grant Applications</i>
66. <i>Volunteer Scheduling Records</i>	66. <i>Volunteer Scheduling Records</i>
67. <i>Real Estate Listing Updates</i>	67. <i>Real Estate Listing Updates</i>
68. <i>Investment Portfolio Rebalancing</i>	68. <i>Investment Portfolio Rebalancing</i>
69. <i>Medical Appointment Schedules</i>	69. <i>Medical Appointment Schedules</i>
70. <i>Travel Itinerary Confirmations</i>	70. <i>Travel Itinerary Confirmations</i>
71. <i>Employment Offer Letters</i>	71. <i>Employment Offer Letters</i>
72. <i>Family Expense Reports</i>	72. <i>Family Expense Reports</i>
73. <i>Home Renovation Cost Estimates</i>	73. <i>Home Renovation Cost Estimates</i>
74. <i>Vehicle Maintenance Schedules</i>	74. <i>Vehicle Maintenance Schedules</i>
75. <i>Insurance Claim Settlements</i>	75. <i>Insurance Claim Settlements</i>
76. <i>Subscription Service Churn Rates</i>	76. <i>Subscription Service Churn Rates</i>
77. <i>Charitable Impact Assessments</i>	77. <i>Charitable Impact Assessments</i>
78. <i>Volunteer Performance Evaluations</i>	78. <i>Volunteer Performance Evaluations</i>
79. <i>Real Estate Market Trends</i>	79. <i>Real Estate Market Trends</i>
80. <i>Investment Risk Assessments</i>	80. <i>Investment Risk Assessments</i>
81. <i>Medical Test Results Analysis</i>	81. <i>Medical Test Results Analysis</i>
82. <i>Travel Itinerary Recommendations</i>	82. <i>Travel Itinerary Recommendations</i>
83. <i>Employment Termination Reasons</i>	83. <i>Employment Termination Reasons</i>
84. <i>Family Financial Planning</i>	84. <i>Family Financial Planning</i>
85. <i>Home Renovation Project Management</i>	85. <i>Home Renovation Project Management</i>
86. <i>Vehicle Recall Resolution Status</i>	86. <i>Vehicle Recall Resolution Status</i>
87. <i>Insurance Policy Renewal Notices</i>	87. <i>Insurance Policy Renewal Notices</i>
88. <i>Subscription Service Feedback</i>	88. <i>Subscription Service Feedback</i>
89. <i>Charitable Grant Disbursement</i>	89. <i>Charitable Grant Disbursement</i>
90. <i>Volunteer Training Completion</i>	90. <i>Volunteer Training Completion</i>
91. <i>Real Estate Listing Status</i>	91. <i>Real Estate Listing Status</i>
92. <i>Investment Portfolio Performance</i>	92. <i>Investment Portfolio Performance</i>
93. <i>Medical Appointment Cancellations</i>	93. <i>Medical Appointment Cancellations</i>
94. <i>Travel Itinerary Modifications</i>	94. <i>Travel Itinerary Modifications</i>
95. <i>Employment Offer Acceptance</i>	95. <i>Employment Offer Acceptance</i>
96. <i>Family Expense Categorization</i>	96. <i>Family Expense Categorization</i>
97. <i>Home Renovation Progress Reports</i>	97. <i>Home Renovation Progress Reports</i>
98. <i>Vehicle Maintenance Cost Analysis</i>	98. <

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6061485"
/tissue_type="melanotic"
/lab_host="DH10 (phage-
/clone.lib="NIH MGC 7"
/note="Organ: skin; Vect
site.2: SalI; Cloned uni
Average insert size 2 kb
Technologies "
```

ORIGIN

Query Match	10.5%;	Score 92;	DB 13;	Length 879;
Best Local Similarity	100.0%;	Prod. No. 1e-154;		
Matches 692;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps 0;				
3964	CGGGAGGAAGCTGCCACCGCCAAAGCAACGGACACCCACACCCCATCCAGCCGAGCCACC	4023		
1	CGGGAGGAAGCTGCCACCGCCAAAGCAACGGACACCCACACCCCATCCAGCCGAGCCACC	60		
4024	CGGAGGCAGCAGATGGCCCATCTCCGCCATCGTGCCTGGCCAGAGCACCAAGCCCGAGTGC	4083		
61	CGGAGGCAGCAGATGGCCCATCTCCGCCATCGTGCCTGGCCAGAGCACCAAGCCCGAGTGC	120		
4084	ATGAGCTCTGGCCCCGCCATCCAGCGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTT	4143		
121	ATGAGCTCTGGCCCCGCCATCCAGCGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTT	180		
4144	AGTCGGCGTCTTAAGGAACGATGCACCAACAATATCTCACCGATTCAACGTAGGACTG	4203		
181	AGTCGGCGTCTTAAGGAACGATGCACCAACAATATCTCACCGATTCAACGTAGGACTG	240		
4204	AACATCGAGGCCACAAAGTGTGCTGTGATACCGTGCACCTTTGAGCCGAGGCA	4263		
241	AACATCGAGGCCACAAAGTGTGCTGTGATACCGTGCACCTTTGAGCCGAGGCA	300		
4264	TCCAAATCTTCGAATGTCAGGTGCATGTGTACCCCAAGTGTCTCCAGTCTCTTCCAGCC	4323		
301	TCCAAATCTTCGAATGTCAGGTGCATGTGTACCCCAAGTGTCTCCAGTCTCTTCCAGCC	360		
4324	ACCTCGCGCTTGCCTGTGTAATATGCCACACACTTCACCGAGGCGCTTCTCGCGTACAAA	4383		

Ds 361 ACCTGGCGCTTGGCTGCTGAATATGCCACACACTTCACCGAGCGCTTCTGCCGTGACAAA 420
Qy 4384 ATGAATCTCCCGAGGTCTCCAGACAAGAGGCCAGAGCAGCTTGGACCTGGAGGGTGG 4443
Ds 421 ATGAATCTCCCGAGGTCTCCAGACAAGAGGCCAGAGCAGCTTGGACCTGGAGGGTGG 480
Qy 4444 ATGAAGTGGCCCGAGGATACAAACAGGAGGACAGAGCGTGGGACAGGAAGTACATTGTC 4503
Ds 481 ATGAAGTGGCCCGAGGATACAAACAGGAGGACAGAGCGTGGGACAGGAAGTACATTGTC 540
Qy 4504 CTGGAGGGATCAAAAGTCTCTCATTTATGACAAATGAAGCCAGAGAGCTGGACAGAGCGCG 4563
Ds 541 CTGGAGGGATCAAAAGTCTCTCATTTATGACAAATGAAGCCAGAGAGCTGGACAGAGCGCG 600
Qy 4564 GTGGAAGAAATTTAGCTGTGCTCTCCGAGCGGGATGTATCTATTGATGTCGCTTGGT 4623
Ds 601 GTGGAAGAAATTTAGCTGTGCTCTCCGAGCGGGATGTATCTATTGATGTCGCTTGGT 660
Qy 4624 GCTTCCGAATCTCGCAATACAGCAAGCAGA 4655
Ds 661 GCTTCCGAATCTCGCAATACAGCAAGCAGA 692

RESULT 7
BC031156
LOCUS BC031156 958 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, similar to citron, clone IMAGE:4976752, mRNA.
ACCESSION BC031156
VERSION BC031156.1 GI:21411076
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 958)
Srausberg.R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-rc@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdcpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 59 Row: j Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: retained intron.

FEATURES
Location/Qualifiers
1..958
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4976752"
/tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI CGAP Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Query Match 10.5%; Score 688.8; DB 11; Length 958;
Best Local Similarity 85.4%; Pred. No. 6.3e-154;
Matches 781; Conservative 0; Mismatches 127; Indels 6; Gaps 1;
Qy 5167 AATTACAGTATCCCTCATTGGAAACCAATAAATCTTACGAATCGACATGAAGCAGTACACG 5226
Ds 40 AAGAAACGATCCCTCATTGGAGCCACAAATTTCTATGATTCGACATGAAGCAGTACACG 99
Qy 5227 CTCGAGGAATTCCTGGATAAGAATGACCATTCCTTGGCACCTGCTGTGTTTGGCGCTCT 5286
Ds 100 CTTGATGAGTTCCTGGACAAGAAGACCATTCCTTGGCACCTGCTGTGTTTGGCGCTCTCG 159
Qy 5287 TCCAAACAGCTTCCTGCTCTCAATCGTCAGTGAAGCGCAGCGGCGAGAGAGTAC 5346
Ds 160 TCCAAACAGCTTCCTGCTCTCAATTTGTCAGCGCAACAGCGCGGCGAGAGAGTAC 219
Qy 5347 TTGCTGTGTTTCCACGAATTTGGAGTCTTCTGGATTTCTTACGGAAGACGTAGCCGACA 5406
Ds 220 CTGCTGTGCTTCCACGAATTTGGGCTCTTCTGATTTCTTACGGAAGACGTAGCCGACA 279
Qy 5407 GACGATCTAAGTGGAGTCTGCTTACCTTTGGCTTTGCCCTACAGAGAACCTTATCTGTTT 5466
Ds 280 GATGATCTTAAAGTGGAGTCTGCTTACCTTGGCTTTGCCCTACAGAGAACCTTATCTGTTT 339
Qy 5467 GTGACCCACTTCAACTCAGTCGAAGTAATTTGAGATCCAGCAGCTCTCTCAGCAGGAGCC 5526
Ds 340 GTGACTCACTTCAACTCTCTGGAAGTCTTTGAGATCCAGGCGCAGATCTCTCAGTGGGAGC 399
Qy 5527 CCTGCCGAGCGTACCTGGACATCCCGAACCCCGCTTACCTGGGCGCTGCAATTCCTCA 5586
Ds 400 CCTGCCGAGCATATCTGGAATTTCCAAACCTCTGCTTACCTGGGCGCGGATTTCTCTCC 459
Qy 5587 GGACGGATTTACTTTGGGCTCTCTATACAGGATAAATTAAGGGTCAATTTCTCTCAAGGGA 5646
Ds 460 GGACGGATTTACTTTGGGCTCTCTATACAGGACAAGTTAAGGGTCAATTTCTCTCAAGGGA 519
Qy 5647 AACCTCTGGAAGGAGTCCGCGACTGAACACACACCGGGCGCCGCTCCACCTCCCGCAGCAGC 5706
Ds 520 AACCTCTGGAAGGAGTCCGCGACTGAGCAGCACCGGGTGCCTCCACCTCCCGCAGCAGC 579
Qy 5707 CCCAAAGCGAGCGCCACCCACCGTACAAAGAGCAGCATCACCAAGCGCGTGGCTCCAGC 5766
Ds 580 CCCAAAGCGAGCGCCACCCACCAATACAAAGAGCAGCATCACCAAGCGCGTGGCTCCAGC 639
Qy 5767 CCAGCGCGCGCCGAAGCGCCCGAGCCACCCCGAGAGCAAGCAGCACACCCACCGCTACC-- 5824
Ds 640 CCGCGCGCACCGGAAGCGCCCGAGCCACCCCGAGAGCAAGCAGCACACCGCACCCTACCGA 699
Qy 5825 ----GCGAGGCGCGGACCGAGCTGCGGAGGGAAGAAGTCTCTGCGCGCGCCCGCTGGAGCGGA 5880
Ds 700 GACAGAGAGGGTCCGACAGAGCTGCGGAGGGAAGAAGTCTCCAGGCGCGCGCTCTGGAGCGG 759
Qy 5881 GAGAAGTCCCGCGCGCGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGCGGTGTTT 5940
Ds 760 GAGAAGTCCCGCGCGCGCGATGCTCAGCATCTAGAGAGAGCGGTCCCGCGGAGAGCTGTTT 819
Qy 5941 GAAGACAGAGCGAGGCGCGCGCTGCTGCGGAGCGCGTGGAGGACCGCGCTGTCCCGAGGTG 6000
Ds 820 GAAGACAGAGCGAGGCGCGCGCTGCTGCGGAGCGAGTGGAGGACCGCGCTGTCCCGAGGT 879
Qy 6001 AACAGGGAAGAGCGGAGAGTCTCTCAAGTCTTACCGTAACTGTCACCTATTAT 6060
Ds 880 AACAGGCTCTGGGACCGAGTCTTCAAGTATAAGTCTCAGCCAAAAAATAAAAAA 939
Qy 6061 GACTGGAATAAAAA 6074
Ds 940 AAAAAAATAAAAAA 953

RESULT 8
BF905370/c

LOCUS BF905370 652 bp mRNA linear EST 18-JAN-2001
DEFINITION IL3-MT0267-261200-410-H07 MT0267 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF905370
VERSION BF905370.1 GI:12296829
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 652)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 2002663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-MT0267-
261200-410-H07&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 624.
FEATURES
Location/Qualifiers
1..652
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0267"
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 9.6%; Score 629.2; DB 10; Length 652;
Best Local Similarity 98.0%; Pred. No. 1e-139;
Matches 637; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 2495 GGAACATGAAGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAATTTACCTGG 2554
DB 652 GGAACATGAAGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAATTTACCTGG 593
QY 2555 AGACACAGCTGGGAAGTTGGAGGCCCAAGAACCGAAAACTGGAGGAGCAGCTGGAGAAGA 2614
DB 592 AGACACAGCTGGGAAGTTGGAGGCCCAAGAACCGAAAACTGGAGGAGCAGCTGGAGAAGA 533
QY 2615 TCAGCCACCAAGACACACAGTGCAGAGATCGGCTCTGGAATCGAGACAGATTCGGGG 2674
DB 532 TCAGCCACCAAGACACACAGTGCAGAGATCGGCTCTGGAATCGAGACAGATTCGGGG 473
QY 2675 AGGTGAGTCTAGACACAGGAGGAGCAAACTGGAGTCAAGGCCCGCAGCTCAGAGAGTAC 2734
DB 472 AGGTGAGTCTAGACACAGGAGGAGCAAACTGGAGTCAAGGCCCGCAGCTCAGAGAGTAC 413
QY 2735 AGCTTCCCTGAGGAGCCGAGTGCAGATTGACAGCCCTGACAGCTGCAGCGCGGCC 2794

DB 412 AGCTCTCCCTGCAGAGCGCGAGTTCACAGTTGACAGCCCTGCAGGCTGCACGGGGGGCCC 353
QY 2795 TGGAGAGCCAGCTTGGCCAGCGGAAGACAGAGCTGGGAAGAGACACAGCAGAGAAGCTGAAG 2854
DB 352 TGGAGAGCCAGCTTGGCCAGCGGAAGACAGAGCTGGGAAGAGACACAGCAGAGAAGCCGAAG 293
QY 2855 AGGAGATCCAGGCACTCAGGSCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTC 2914
DB 292 AGGAGATCCAGGCACTCAGGSCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTC 233
QY 2915 GTAACAGCTGTACTGTATATCAGACAGCTGGAGGAGCAGCTAAACCACTGACCGAGGACA 2974
DB 232 GTAACAGCTGTACTGTATATCAGACAGCTGGAGGAGCAGCTAAACCACTGACCGAGGACA 173
QY 2975 ACGCTGAACCTCAACAAACCAAACTTCTACTTGTCTCAACAACTCGATGAGGCTTCTGCG 3034
DB 172 ACGCTGAACCTCAACAAACCAAACTTCTACTTGTCTCAACAACTCGATGAGGCTTCTGCG 113
QY 3035 CCAACGACGAGATTGTACAACTCGAACTGGAAGTGAAGTGGACCTCTCCGCGGAGATCAGG 3094
DB 112 CCAACGACGAGATTGTACAACTCGAACTGGAAGTGGACCTCTCCGCGGAGATCAGG 53
QY 3095 AACGAGAGATGCAGCTTACCAGCCAGAGCAACAGCATGGAGGCTCTGAAG 3144
DB 52 AACGAGAGATGCAGCTTACCAGCCAGAGCAACAGCATGGAGGCTTGAACAAG 3
RESULT 9
LOCUS BG976452 881 bp mRNA linear EST 12-JUN-2001
DEFINITION 602846269P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4976752 5',
mRNA sequence.
ACCESSION BG976452
VERSION BG976452.1 GI:14364089
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10970 row: j column: 17
High quality sequence stop: 841.
FEATURES
Location/Qualifiers
1..881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4976752"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
ORIGIN
Query Match 9.2%; Score 606.6; DB 12; Length 881;

Db 301 CAGGTGGTGGAGTGGCGGTGAAGAACACAAAGCGCGGAGATTCTTGCTCTGCAGCAGGCT 360
QY 3415 CTCAAAGACGAGAAGCTGAAGCCGAGAGCCCTCTCTGACAAAGCTCAATGACCTGAGAG 3474
Db 361 CTCAGGACGACGAGCTCAAGCCGAGAGCCTGTGGACAAAGCTCAACGACCTGAGAG 420
QY 3475 AAGCATGCTATGCTTGAATGAATGCCGAAAGCTTACACAGAACTGAGAGCTGAAACGA 3534
Db 421 AAGCAGCCCATGCTGGAGATGAAGCCCGGAGCTTACAGCAGAACTAGAGACAGAGCGG 480
QY 3535 GAGCTCAACACAGGCTCTGAGAGCAAGCAAGCCAAATTACACAGCAGATGACCTGCGAG 3594
Db 481 GAGCTCAACACAGGCTCTGAGAGCAAGCAAGCCAAATTACACAGCAGATGACCTGCGAG 540
QY 3595 AAAAAATCACTTTCCGCTGACTCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTA 3654
Db 541 AAGAACCATCTCTCAGACTGACGCAAGGCTGACGAGGCGCTGACCGCGGATCTG 600
QY 3655 CTCGAGACGAGAAGAGTCACTTGGAGTATCAGCTGAGTGGAAACATTCAGGTTCTCTATCT 3714
Db 601 CTCGAGACGAGAAGGAGGACCTGAGTACCAGCTGGAAACATTCAGGTTCTCTACTCT 660
QY 3715 CATGAAAGAGTGAAATGAAGGCACTATTCTCAACAAACCAAACTCATTTGATTTCTG 3774
Db 661 CACGAGAAAGTGAAATGAAGGCACTATTCTCAGCANA-CANACTCATTTGATTTCTG 719
QY 3775 CAGCCAAATGACCAACTGCTAAAA 3802
Db 720 CAAG-CANAATGACCGCTGCTAAAA 746

RESULT 11

LOCUS BG912161 843 bp mRNA linear EST 05-JUN-2001
DEFINITION 602812833F1 NCI_CGAP_Brn67 Homo sapiens cdna clone IMAGE:4944657
5', mRNA sequence.
ACCESSION BG912161
VERSION BG912161.1 GI:14292637
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1. (bases 1 to 843)
TITLE NIH-MGC <http://mgi.nci.nih.gov/>
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10890 row: a column: 10
High quality sequence stop: 778.

FEATURES

source
1..843
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4944657"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn67"
/note="Organ: Brain; Vector: pCMV-SPOrt6; Site 1: Nct1;
Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 8.9%; Score 584.6; DB 12; Length 843;
Best Local Similarity 92.8%; Pred. No. 6.9e-129;
Matches 669; Conservative 0; Mismatches 44; Indels 8; Gaps 5;
QY 4654 GAAAAAGCAGAGCTGATGCTAACTCTGCTGGAACCTCCCTGCTGAACTGGAAGGTGAT 4713
Db 122 GAAAAAGCAGAGCTGATGCTAACTCTGCTGGAACCTCCCTGCTGAACTGGAAGGTGAT 181
QY 4714 GACCGTCTAGACATGAATCTGACGCTCCCTTCAGTGACAGGTGGTGGTGGGCACC 4773
Db 182 GACCGTCTAGACATGAATCTGACGCTCCCTTCAGTGACAGGTGGTGGTGGGCACC 241
QY 4774 GAGAAAGGCTCTACGCCCTGAATGCTTTGAAAAAATCTCCCTAACCCATGTCCCGAGAAAT 4833
Db 242 GAGAAAGGCTCTACGCCCTGAATGCTTTGAAAAAATCTCCCTAACCCATGTCCCGAGAAAT 301
QY 4834 GGACGCTCTCCAAATTTATATATCAAGCACTCGAGAGCTGAGAGCTACTCATGATAGAGGA 4893
Db 302 GGACGCTCTCCAAATTTATATATCAAGCACTCGAGAGCTGAGAGCTACTCATGATAGAGGA 361
QY 4894 GAAGAGGGGCACTGTGCTTGTGGAGCTGAAAGAACTGAAACAGTCCCTGGCCCACTCC 4953
Db 362 GAAGAGGGGCACTGTGCTTGTGGAGCTGAAAGAACTGAAACAGTCCCTGGCCCACTCC 421
QY 4954 CACCTGCTGCCGAGCCCGACATCTCACCCAACTTTTGAAGCTGTCAAGGGCTGCCAC 5013
Db 422 CACCTGCTGCCGAGCCCGACATCTCACCCAACTTTTGAAGCTGTCAAGGGCTGCCAC 481
QY 5014 TTGTTTGGGCGAGCAAGATTGAAACGGGCTCTGCATCTGTGCAGCCATGCCCGACAAA 5073
Db 482 TTGTTTGGGCGAGCAAGATTGAAACGGGCTCTGCATCTGTGCAGCCATGCCCGACAAA 541
QY 5074 GTCGTCATTTCCGCTCAACGAAAACTCAGCAAAATCTGCATCCCGAAAGAGATAGAG 5133
Db 542 GTCGTCATTTCCGCTCAACGAAAACTCAGCAAAATCTGCATCCCGAAAGAGATAGAG 601
QY 5134 ACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTTACAGTATCTCTCAAT- 5192
Db 602 ACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTTACAGTATCTCTCAATTTGGAACCA 661
QY 5193 TAAATTTACG- AATCGACATGAA- GCAGTACACGCTCGAGGAAT- -CCTGGATAAGA 5248
Db 662 TAAATTTACGAAAAATCGACATGAAGCGAGTACACGCTCGAGGAATTTCTTGGATAACGA 721
QY 5249 ATGACCATTTCTTGGCACCTGCTGTG- -TTTGGCGGCTCTTCCACAGTCTCCCTGCT 5305
Db 722 ATGACCATTTCTTGGCACCTGCTGTG- -TTTGGCGGCTCTTCCACAGTCTCCCTGCT 781
QY 5306 CAATCGTGCAGGTGAACAGCGCGAGGCGAGGAGAGTACTTGTCTGTCTTCCACGAAT 5365
Db 782 AATCGTGCAGGTGAACAGCGCGAGGCGAGGAGAGTATTGCTGTCTTCCCGGACT 841
QY 5366 T 5366
Db 842 T 842

RESULT 12

LOCUS CF744580 647 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-GVO-clt-n-23-0-UI.r1 NIH BMAP_GVO Mus musculus cdna clone
IMAGE:30617710 5', mRNA sequence.
ACCESSION CF744580
VERSION CF744580.1 GI:37640920
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS 1. (bases 1 to 647)
TITLE NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES
source

Seq primer: pYX-5.
Location/Qualifiers
1..647
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30617710"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 8.7%; Score 573.6; DB 14; Length 647;
Best Local Similarity 93.8%; Pred. No. 2.5e-126;
Matches 608; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 1149 CACCTCAAGTCGACGATGACACCTCCCAATTTTGATGAACAGAGAAGAAATTCGTGGGT 1208
DB 1 CACCTCAAGTCGACGATGACACCTCCCAATTTTGATGAACAGAGAAGAAATTCGTGGGT 60
QY 1209 TTCATCCTCTCCGTGCCAGCTGAGCCCTCAGCCTTCCTGGGTGAAGAACTCCGCTTTGT 1268
DB 61 TTCATCCTCTCTGTGCCAGCTGAGCCCTCGGCTTCTCAGGCAAGAGCTGCGCTTGT 120
QY 1269 GGGGTTTCTAGCAAGCACTGGGATTTCTGTAGTCTGAGTCTGTGTCTCGG 1328
DB 121 GGGATTTCTGACAGCAAGCACTGGGATTTCTGTAGTCTGAGTCTGTGTCTCGG 180
QY 1329 TCTGGACTCCCTCCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGA 1388
DB 181 TCTGGACTCCCTCCCAAGCTTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGA 240
QY 1389 GCTACAGACTCTCAGGACGAGTGTCAAGATGGAGCGGAAATGACCGGTACATCG 1448
DB 241 GCTCAGACTCTCCAGCAAGTGTCAAGATGGAGCGGAAATGACCGGTACATCG 300
QY 1449 GAGAGTGTCTCAGAGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTGA 1508
DB 301 CAGAGTCTCAGAGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTGA 360
QY 1509 GACTCAGAGTCCCTCTCTGGAGCAGACCTTGTACTACATCAGCAAGATGAGTAGTCT 1568
DB 361 GACTCAGAGTCCCTCTCTGGAGCAGACCTTGTACTACATCAGCAAGATGAGTAGTCT 420
QY 1569 AAAGCGAAGTTTGGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGCA 1628

DB 421 AAAGCGAAGTTTGGAGCAAGCAGGATGGAGGTCTCCAGGAGATGACAAAGCTCTGCA 480
QY 1629 GCTTCTCATGATATCAGAGAGCAGCGGAACTCCAAAGAAATCAAAGAGAGGAGTA 1688
DB 481 GCTTCTCCACACATCCGAGAGCAGCGGAACTCCAGGATCAAGAGAGGAGTA 540
QY 1689 CCAGGCTCAAGTCCGAAGAAATAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTC 1748
DB 541 CCAGGCTCAGGTGAGGAGATGAGCTGATGATGAATCAGTGGAGAGGA-CTGTGTCTC 599
QY 1749 AGCAAGAGAGCGGATGATCTCTACGAATCTGAGCTGAGAGAGTCTCG 1796
DB 600 AGCCCGCAGACGCGAGGATCTCTACGAGTCTGAGCTGAGGAGTCTCG 647

RESULT 13
CF723360

LOCUS 730 bp mRNA linear EST 09-OCT-2003
DEFINITION UI-M-GVO-cjh-j-20-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
IMAGE:30546187 5', mRNA sequence.

ACCESSION

CF723360

VERSION

CF723360.1

KEYWORDS

GI:37597528

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 730)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers
1..730
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30546187"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 8.7%; Score 572.8; DB 14; Length 730;
Best Local Similarity 88.7%; Pred. No. 4.3e-126;

LOCUS B0421272 894 bp mRNA linear EST 23-MAY-2002
DEFINITION AGENCOURT 7761930 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6018187
5', mRNA sequence.
ACCESSION B0421272
VERSION B0421272.1 GI:21116587
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 894)
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: egapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LHAM13218 row: k column: 20
High quality sequence stop: 720.
Location/Qualifiers
1..894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6018187"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

FEATURES
source

ORIGIN
Query Match 8.7%; Score 570; DB 13; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6005 AGGAGAGAGGCGAGAGTCTCTCAAGTTTTCACGGTTAACACTGTACCTATTATGACT 6064
DB 570 AGGAGAGAGGCGAGAGTCTCTCAAGTTTTCACGGTTAACACTGTACCTATTATGACT 511
QY 6065 GGAATATAAAGCTGGACCACTGCCAGCTAACTGGTTCAGTCTGAGGATCATCCAGCTGA 6124
DB 510 GGAATATAAAGCTGGACCACTGCCAGCTAACTGGTTCAGTCTGAGGATCATCCAGCTGA 451
QY 6125 ATGGAGAAATCCGGCAGCAGGTTGAAAAGTCTCTTCTGAGAACAGATTTATGCTGAGCAG 6184
DB 450 ATGGAGAAATCCGGCAGCAGGTTGAAAAGTCTCTTCTGAGAACAGATTTATGCTGAGCAG 391
QY 6185 AGTTTCATGTGACTCTAGAGCTGGTGAATTAATAAATGGCTTAAGCTGACAGCCAGC 6244
DB 390 AGTTTCATGTGACTCTAGAGCTGGTGAATTAATAAATGGCTTAAGCTGACAGCCAGC 331
QY 6245 CACCTCTGCTTACAAAAGAGTACTTAGTGCACATGACTGTAAAGAAACAATTGTAACCC 6304
DB 330 CACCTCTGCTTACAAAAGAGTACTTAGTGCACATGACTGTAAAGAAACAATTGTAACCC 271
QY 6305 TCATCTAGAAATCAGAAAGCTCTTAATTTCTATAGAAATGACACCTCCCTGGAGCCGAGA 6364
DB 270 TCATCTAGAAATCAGAAAGCTCTTAATTTCTATAGAAATGACACCTCCCTGGAGCCGAGA 211
QY 6365 GACAATCTGTTGTGATTTTGAAGGAGCAGGCAAGACCAACACTGTATTTAGTTCCATAGC 6424
DB 210 GACAATCTGTTGTGATTTTGAAGGAGCAGGCAAGACCAACACTGTATTTAGTTCCATAGC 151
QY 6425 CAGGCCTCAACAGGACCAAGTGGCTTAAACACACACAGATGACTGGAAATGATGT 6484

Db 150 CAGGCCTCAACAGGACCAAGTGGCTTAAACACACAGATGACTGGAAATGATGT 91
QY 6485 GTGGCCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGAGTTAGCATTCAATTTTGG 6544
Db 90 GTGGCCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGAGTTAGCATTCAATTTTGG 31
QY 6545 CTTAAGAAAAATCGAGAATGTAGGTTTAGA 6574
Db 30 CTTAAGAAAAATCGAGAATGTAGGTTTAGA 1

Search completed: July 4, 2004, 07:39:38
Job time : 14343 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:07:00 ; Search time 104 Seconds
(without alignment)
5577.596 Million cell updates/sec

Title: US-10-017-216-2
Perfect score: 10490
Sequence: 1 MLKFYGYARNPLDAGAEPI.....QLNGSIROQVEKSVLRTDYC 2053

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_28Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	10490	100.0	2053	5	AEE24079	Human MDP	AEE24079 Human MDP
2	10032.5	95.6	2054	5	AEE24150	Human kin	AEE24150 Human kin
3	10032.5	95.6	2054	6	AAC26959	Human CRI	AAC26959 Human CRI
4	10022.5	95.5	2054	5	ABB81927	Human kin	ABB81927 Human kin
5	10012	95.4	2053	4	AU03501	Human pro	AU03501 Human pro
6	10007	95.4	2066	5	ABG78362	Human pro	ABG78362 Human pro
7	10007	95.4	2066	6	ADA05654	Human NOV	ADA05654 Human NOV
8	10005	95.4	2053	5	ABG78363	RHO/RAC-i	ABG78363 RHO/RAC-i
9	10005	95.4	2053	6	ADA05642	Human NOV	ADA05642 Human NOV
10	9656	92.0	2055	6	ABP97683	Polypepti	ABP97683 Polypepti
11	9656	92.0	2055	6	AAC26960	Human CRI	AAC26960 Human CRI
12	9487.5	90.4	1958	5	ABB81928	Human kin	ABB81928 Human kin
13	7521	71.7	1619	7	ADD48584	Rat Prote	ADD48584 Rat Prote
14	7521	71.7	1619	7	ADD46616	Rat Prote	ADD46616 Rat Prote
15	6223.5	59.3	1286	3	AAB43359	Human ORF	AAB43359 Human ORF
16	6223.5	59.3	1286	6	AAC26961	Human CRI	AAC26961 Human CRI
17	6223.5	59.3	1286	7	ADD48586	Human Pro	ADD48586 Human Pro
18	6223.5	59.3	1286	7	ADD46618	Human Pro	ADD46618 Human Pro
19	4710.5	44.9	999	4	ABB11117	Human RHO	ABB11117 Human RHO
20	4519.5	43.1	940	7	ADD89967	Human can	ADD89967 Human can
21	4174.5	39.8	910	4	AU31443	Novel hum	AU31443 Novel hum
22	4101	39.1	847	6	ADA05648	Human NOV	ADA05648 Human NOV
23	4005.5	38.2	832	6	ADA05646	Human NOV	ADA05646 Human NOV
24	3976.5	37.9	853	7	ADG09957	Novel pro	ADG09957 Novel pro
25	3155	30.1	623	6	ADA05644	Human NOV	ADA05644 Human NOV

26	2910	27.7	623	6	ADA05650	Human NOV	Ada05650 Human NOV
27	2892.5	27.6	638	6	ADA05652	Human NOV	Ada05652 Human NOV
28	2433	23.2	495	6	ABP97681	Amino aci	ABP97681 Amino aci
29	2433	23.2	497	6	ABP97687	Novel hum	ABP97687 Novel hum
30	2430	23.2	497	6	ABU10126	Novel hum	ABU10126 Novel hum
31	2430	23.2	497	7	AEE39504	Human kin	AEE39504 Human kin
32	2425	23.1	497	5	AEE16261	Human kin	AEE16261 Human kin
33	2417	23.0	497	6	ABU10127	Variant n	ABU10127 Variant n
34	2185.5	20.6	494	6	ABP97682	Polypepti	ABP97682 Polypepti
35	1620	15.4	349	4	ABG15566	Novel hum	ABG15566 Novel hum
36	1446	13.8	1770	5	AEE25099	Human kin	AEE25099 Human kin
37	1421.5	13.6	1719	5	AEE21707	Human PKI	AEE21707 Human PKI
38	1400	13.3	1738	7	ADE47738	Human NOV	ADE47738 Human NOV
39	1399.5	13.3	1664	7	ADE47740	Human NOV	ADE47740 Human NOV
40	1398.5	13.3	1797	4	ABG13880	Novel hum	ABG13880 Novel hum
41	1390	13.3	1728	3	ABG42069	Human ORF	ABG42069 Human ORF
42	1390	13.3	1728	6	AAB38363	Human ser	AAB38363 Human ser
43	1382.5	13.2	1711	5	AAU97065	CDC42-bin	AAU97065 CDC42-bin
44	1382.5	13.2	1711	6	AEE38364	Human ser	AEE38364 Human ser
45	1329	12.7	1252	4	ABB66357	Drosophil	ABB66357 Drosophil

ALIGNMENTS

RESULT 1
AAE24079
ID AAE24079 standard; protein; 2053 AA.
XX
AC AAE24079;
XX
DT 04-OCT-2002 (first entry)
XX
DE Human MDPK protein.
XX
KW Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;
KW tumorigenesis; tumour growth; tumour metastasis; viral infection;
KW skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;
KW immune disorder; neoplastic disorder; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 53..303
FT Domain /note= "pkinase domain"
FT Domain 97..360
FT Domain /note= "pkinase domain"
FT Region 195..210
FT Region /note= "Antigenic epitope"
FT Region 217..229
FT Region /note= "Serine/Threonine protein kinase active site signature"
FT Region 455..475
FT Domain /note= "Antigenic epitope"
FT Domain 1568..1865
FT Domain /note= "CNH domain"
WO200234896-A2.
02-MAY-2002.
23-OCT-2001; 2001WO-US050636.
23-OCT-2000; 2000US-0242429P.
(MILL-) MILLENNIUM PHARM INC.
Kapeller-Libermann R;
WPI; 2002-479720/51.
N-PSDB; AAD39191.
Human myotonic dystrophy type protein kinase polypeptide and

PT polynucleotide useful for prognosticating, diagnosing, preventing or
 PT inhibiting tumorigenesis, tumor growth, tumor metastasis and viral
 XX infection.

XX Claim 8; Fig 3; 148pp; English.

XX The invention relates to human myotonic dystrophy type protein kinase
 CC (MDPK) polypeptides designated as 13245 and nucleic acid molecules
 CC encoding such polypeptides. 13245 molecules are used to develop
 CC diagnostic and therapeutic agents for prognosticating, diagnosing,
 CC preventing, inhibiting, alleviating or curing MDPK-related disorders.
 CC Polypeptides of the invention are used to develop diagnostic and
 CC therapeutic agents for 13245-mediated or related disorders such as
 CC tumorigenesis, tumor growth, tumor metastasis, viral infection of a
 CC cell, skeletal muscle disorders (e.g. muscular and myotonic dystrophies),
 CC immune disorders and neoplastic disorders. The invention is also used in
 CC gene therapy. The present sequence is human MDPK protein
 XX

SQ Sequence 2053 AA;

Query Match 100.0%; Score 10490; DB 5; Length 2053;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLKFKYGARNPLDAGAAEPIASRASRLNLFQKPPFMTQQQMSPLSREGILDALFVLFE	60
DB	1	MLKFKYGARNPLDAGAAEPIASRASRLNLFQKPPFMTQQQMSPLSREGILDALFVLFE	60
QY	61	ECQPALMKIKHVSFVRKYSDTIAELOELQPSAKDFEVRSLVGCCHFAEVOVREKATG	120
DB	61	ECQPALMKIKHVSFVRKYSDTIAELOELQPSAKDFEVRSLVGCCHFAEVOVREKATG	120
QY	121	DIYAMKVMKKALLAQBOVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLYLMEEYQPGG	180
DB	121	DIYAMKVMKKALLAQBOVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLYLMEEYQPGG	180
QY	181	DLISLNLRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVD	240
DB	181	DLISLNLRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVD	240
QY	241	GSAAKMNSKVMNAKPIGTPDYMAPEVLTVMNGDGKTYGLDCDWSVGLVAYEMVYGR	300
DB	241	GSAAKMNSKVMNAKPIGTPDYMAPEVLTVMNGDGKTYGLDCDWSVGLVAYEMVYGR	300
QY	301	SPFAEGTSARTFNNINNFORFLKPPDDPKVSSDFLDLIQSLCGQKERLKFEGLCCHPFF	360
DB	301	SPFAEGTSARTFNNINNFORFLKPPDDPKVSSDFLDLIQSLCGQKERLKFEGLCCHPFF	360
QY	361	SKIDMNNIRNSPPFPVTLKSDDDTSNDFEPEKNSVWSSPCQLSPSGSGBELPVGFS	420
DB	361	SKIDMNNIRNSPPFPVTLKSDDDTSNDFEPEKNSVWSSPCQLSPSGSGBELPVGFS	420
QY	421	YSKALGILGRSESVSGLDSPAKTSNWEKKLIATKSKELQDSQDKCHKMEQEMTRLHRRVS	480
DB	421	YSKALGILGRSESVSGLDSPAKTSNWEKKLIATKSKELQDSQDKCHKMEQEMTRLHRRVS	480
QY	481	EVPAVLSQKEVELKASQTSRLLEQDLATYITFCSLSKSLKSLQARNEVSEDDKALQLLH	540
DB	481	EVPAVLSQKEVELKASQTSRLLEQDLATYITFCSLSKSLKSLQARNEVSEDDKALQLLH	540
QY	541	DIREQRKLQEIKEQYQAOVEEMRLMMQLEEDLVSAARRRDLVSELSRESRLAAEEFK	600
DB	541	DIREQRKLQEIKEQYQAOVEEMRLMMQLEEDLVSAARRRDLVSELSRESRLAAEEFK	600
QY	601	RKATECQHLKAKQCKQKPVGVYAKLEKINAEQQLKIQELOKLEKAAKAEARELEKL	660
DB	601	RKATECQHLKAKQCKQKPVGVYAKLEKINAEQQLKIQELOKLEKAAKAEARELEKL	660
QY	661	QNRDSSGIRKKLVAEAEERHSLKVKLEMTERRERNLKDDIOTKQQQIQQADKIL	720
DB	661	QNRDSSGIRKKLVAEAEERHSLKVKLEMTERRERNLKDDIOTKQQQIQQADKIL	720
QY	721	ELBEKHREAQVSAQHLVHLKQEQHYEKKIKVLDNQIKKDLADKETLENMQRHEEAAH	780

DB	721	ELBEKHREAQVSAQHLVHLKQEQHYEKKIKVLDNQIKKDLADKETLENMQRHEEAAH	780
QY	781	EKGKILSEQAMINAMDSKIRSLFORIVELSEANKLAANSSLTORMNKAQEEMISBLRQ	840
DB	781	EKGKILSEQAMINAMDSKIRSLFORIVELSEANKLAANSSLTORMNKAQEEMISBLRQ	840
QY	841	QKFVLETCAGKLAQNEKLEBEQLEKISHQSHDSKNRLLLETLRLREVSLHEHQKLELKR	900
DB	841	QKFVLETCAGKLAQNEKLEBEQLEKISHQSHDSKNRLLLETLRLREVSLHEHQKLELKR	900
QY	901	QLTELOLSLOERESQLTALQAARAALESQLRQAKTELETTAAEETTAALTAHRDEIQ	960
DB	901	QLTELOLSLOERESQLTALQAARAALESQLRQAKTELETTAAEETTAALTAHRDEIQ	960
QY	961	KFDALRNSCTVITLBEOLNQLTDEANLNQNFYLSKQLEASGANDETVQLRSEVDHL	1020
DB	961	KFDALRNSCTVITLBEOLNQLTDEANLNQNFYLSKQLEASGANDETVQLRSEVDHL	1020
QY	1021	RREITEREMOLTSQKOTMEALKTTCTMLEBQVMDLEALNDELLEKEREQWEANRSVLGDEK	1080
DB	1021	RREITEREMOLTSQKOTMEALKTTCTMLEBQVMDLEALNDELLEKEREQWEANRSVLGDEK	1080
QY	1081	SQFCRVRELOQMLDTEKQSRADQIRITESQVVELAVKEHKAETIALQOALKEQKLLKA	1140
DB	1081	SQFCRVRELOQMLDTEKQSRADQIRITESQVVELAVKEHKAETIALQOALKEQKLLKA	1140
QY	1141	ESLSKDLNLEKXHAMLENNARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRLT	1200
DB	1141	ESLSKDLNLEKXHAMLENNARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRLT	1200
QY	1201	QGLQALDRADLLKTERSDLEYOLENIQVLYSHKVMMEGTISQOQTKLIDFLQAKMDQPA	1260
DB	1201	QGLQALDRADLLKTERSDLEYOLENIQVLYSHKVMMEGTISQOQTKLIDFLQAKMDQPA	1260
QY	1261	KKKKGLFSRRKEDPALPTQVPLQVNEUKLALKEKARCABEALOKTRIELSAREEAA	1320
DB	1261	KKKKGLFSRRKEDPALPTQVPLQVNEUKLALKEKARCABEALOKTRIELSAREEAA	1320
QY	1321	HRKATDHPHSTPATARQQIAMSALVRSPEHQSAMSLSLAPPSRRKESSTPEFSRRLK	1380
DB	1321	HRKATDHPHSTPATARQQIAMSALVRSPEHQSAMSLSLAPPSRRKESSTPEFSRRLK	1380
QY	1381	ERMHNIPHRFNVLNMRATKCAVCLDTVHFGQAQSKLCEQVWCHPKCSTCLPATCGLP	1440
DB	1381	ERMHNIPHRFNVLNMRATKCAVCLDTVHFGQAQSKLCEQVWCHPKCSTCLPATCGLP	1440
QY	1441	AEYATHTEAFCDKKNMSPGLQTKPESSSLHLEGMKVPNNKRGQGGWDRKYIVLEGSK	1500
DB	1441	AEYATHTEAFCDKKNMSPGLQTKPESSSLHLEGMKVPNNKRGQGGWDRKYIVLEGSK	1500
QY	1501	VLIYDNEAREAGQRPVEEFELCLPDGQVSIHGAVGASELANATAKAEKADAKLGNLSL	1560
DB	1501	VLIYDNEAREAGQRPVEEFELCLPDGQVSIHGAVGASELANATAKAEKADAKLGNLSL	1560
QY	1561	KLEGGDLDMNCTLPFSDQVVLVGTTEGLVALNVLKNSLTHVPGIGAVFQIYIKOLEKL	1620
DB	1561	KLEGGDLDMNCTLPFSDQVVLVGTTEGLVALNVLKNSLTHVPGIGAVFQIYIKOLEKL	1620
QY	1621	LMTAGEBREALCLVDVKVKQSLAQSHLPAQDPISPNIFEAVKCHLPGAGKIENGLCICA	1680
DB	1621	LMTAGEBREALCLVDVKVKQSLAQSHLPAQDPISPNIFEAVKCHLPGAGKIENGLCICA	1680
QY	1681	AMPSKVVILRYNENLSKYCIKBTETSEPCSCITFTNYSILIGTNKPYEDMKQYTLBEF	1740
DB	1681	AMPSKVVILRYNENLSKYCIKBTETSEPCSCITFTNYSILIGTNKPYEDMKQYTLBEF	1740
QY	1741	LDKNDHSLAPAVFAASSNSPFVSIVQNSAGQREPEVILLCFHEFGVFDVSGRSRTDILK	1800
DB	1741	LDKNDHSLAPAVFAASSNSPFVSIVQNSAGQREPEVILLCFHEFGVFDVSGRSRTDILK	1800
QY	1801	WSRLPLAFAYREPVLFTYHNSLSEVIEIOARSAGTTPARAVLDIPNPRYGPATSSGAIY	1860


```

Db      1801  WRLPLAFAYREFFLVTHFNLSLEVIEIOARSSAGTAPARAYLDIPNRYLQPAISGAIY 1860
QY      1861  LASSYQDLRVICCKGNLVKESGTEHRRGPSTSRSSPNKRGPTTYNEHTKEVASSPAPP 1920
Db      1861  LASSYQDLRVICCKGNLVKESGTEHRRGPSTSRSSPNKRGPTTYNEHTKEVASSPAPP 1920
QY      1921  EGPSPREPSTPHRYREGTELRDDKSGPRLERKSPGRLMSTRERSPGKLPEDSSRG 1980
Db      1921  EGPSPREPSTPHRYREGTELRDDKSGPRLERKSPGRLMSTRERSPGKLPEDSSRG 1980
QY      1981  RLPAGAVRTPLSQVNGKRGQSASQVFTVTNTVYDWNKKLDNLNPNWVSLRIQLNGEIR 2040
Db      1981  RLPAGAVRTPLSQVNGKRGQSASQVFTVTNTVYDWNKKLDNLNPNWVSLRIQLNGEIR 2040
QY      2041  QQVEKSVLRDYC 2053
Db      2041  QQVEKSVLRDYC 2053

RESULT 2
AAE24150
ID      AAE24150 standard; protein; 2054 AA.
AC      AAE24150;
DT      23-SEP-2002 (first entry)
XX      Human kinase (PKIN)-21 protein.
DE      Human kinase (PKIN)-21 protein.
XX      Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
KW      acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
KW      asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
KW      development; hepatitis; cardiovascular; hypertension; drug screening;
KW      myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
KW      fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
KW      hypercholesterolemia; obesity; gene therapy; cytostatic; anti-hiv;
KW      neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
XX      Homo sapiens.
OS      Homo sapiens.
XX      Key
FH      Location/Qualifiers
FT      97..360
FT      Domain
FT      /note= "Eukaryotic protein kinase domain"
FT      98..241
FT      Domain
FT      /note= "Protein kinase domain"
FT      99..349
FT      Domain
FT      /note= "Protein kinase domain"
FT      101..241
FT      Domain
FT      /note= "Protein kinase domain"
FT      102..241
FT      Domain
FT      /note= "Protein kinase domain"
FT      249..349
FT      Domain
FT      /note= "Protein kinase domain"
FT      258..445
FT      Domain
FT      /note= "Protein kinase domain"
FT      258..349
FT      Domain
FT      /note= "Protein kinase domain"
FT      534..542
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      544..875
FT      Region
FT      /note= "Leucine zipper pattern"
FT      891..933
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      964..975
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      991..1012
FT      Region
FT      /note= "Leucine zipper pattern"
FT      1015..1067
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      1057..1078
FT      Region
FT      /note= "Leucine zipper pattern"
FT      1159..1180
FT      Region

```

```

FT      Domain
FT      /note= "Leucine zipper pattern"
FT      1217..1255
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      1388..1434
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      1390..1438
FT      Binding-site
FT      /note= "Phorbol esters/diacylglycerol binding site"
FT      1403..1466
FT      Binding-site
FT      /note= "Phorbol esters/diacylglycerol binding site"
FT      1471..1590
FT      Domain
FT      /note= "PH domain"
FT      1619..1916
FT      Domain
FT      /note= "CNH domain"
FT      1759..1802
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      1819..1831
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
FT      1851..1880
FT      Domain
FT      /note= "Domain found in NIKI-like kinase"
XX      WO200233099-A2.
XX      25-APR-2002.
XX      20-OCT-2001; 2001WO-US047728.
XX      20-OCT-2000; 2000US-0242410P.
XX      27-OCT-2000; 2000US-0244068P.
XX      03-NOV-2000; 2000US-0245708P.
XX      09-NOV-2000; 2000US-024672P.
XX      16-NOV-2000; 2000US-0249565P.
XX      22-NOV-2000; 2000US-0252730P.
XX      01-DEC-2000; 2000US-0250807P.
XX      (INCY-) INCYTE GENOMICS INC.
XX      Gururajan R, Baughn MR, Wallia NK, Elliott VS, Xu Y, Arvizu C;
XX      Yao MG, Rankumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;
XX      Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;
XX      Recipon SA, Lu DAW, Borowsky ML, Thornton M, Swarnaker A;
XX      Thangavelu K, Khan FA, Ison CH;
XX      WPI; 2002-454603/48.
XX      N-PSDB; AAD38864.
XX      New human kinase polypeptide, for diagnosing, preventing and treating
XX      cancer, immune system disorders, growth and development disorders,
XX      cardiovascular disorders and lipid disorders.
XX      Claim 1; Page 177-182; 210pp; English.
XX      The invention relates human kinases (PKIN) and their corresponding
XX      nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
XX      treating and preventing cancer, an immune system disorder (e.g., acquired
XX      immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
XX      atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
XX      growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
XX      cardiovascular disorder (e.g., hypertension, myocardial infarction,
XX      Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
XX      Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,
XX      hyperlipidaemia, obesity), and for assessing the effects of exogenous
XX      compounds. Anti-PKIN antibody is useful in a diagnostic test for a
XX      condition or a disease associated with the expression of PKIN in a
XX      biological sample. A composition comprising PKIN or an agonist or
XX      antagonist of PKIN is useful for treating a disease or condition
XX      associated with decreased or increased expression of functional PKIN.
XX      PKIN is useful in a number of drug screening techniques and to analyse
XX      the proteome of a tissue or cell type. PKIN DNA is useful for creating
XX      knockin humanised animals or transgenic animals to model human diseases,
XX      and in somatic or germline gene therapy. The present sequence is human
XX      PKIN protein
XX      Sequence 2054 AA;
SQ

```

Query Match	95.6%;	Score	10032.5;	DB 5;	Length	2054;
Best Local Similarity	95.8%;	Pred. No.	0;			
Matches 1981;	Conservative	2;	Mismatches	4;	Indels	81; Gaps 3;
Qy	1	MLRFKYGAR	NPLDAGAAEPIAS	RASRLNLPFOQKPPFMTQQQVSP	LSRBGILDALFVLPE	60
Db	1	MLRFKYGAR	NPLDAGAAEPIAS	RASRLNLPFOQKPPFMTQQQVSP	LSRBGILDALFVLPE	60
Qy	61	ECSQPALMKIKHVS	NFVKYSDTIAELQELQPSAKDFE	VRSVLGCGGHFAEVOVVRKATG	120	
Db	61	ECSQPALMKIKHVS	NFVKYSDTIAELQELQPSAKDFE	VRSVLGCGGHFAEVOVVRKATG	120	
Qy	121	DIYAMKVMKKALLA	QOVQVFFBEERNILSRSTSPWIPOLQ	YAFQDKNHLHLYMEETQPGG	180	
Db	121	DIYAMKVMKKALLA	QOVQVFFBEERNILSRSTSPWIPOLQ	YAFQDKNHLHLYMEETQPGG	180	
Qy	181	DLISLNLNRYEDQ	LDENLIQVLAELILAVHSVHLMGVVHR	DIKPENILVDRTGHIKLVD	240	
Db	181	DLISLNLNRYEDQ	LDENLIQVLAELILAVHSVHLMGVVHR	DIKPENILVDRTGHIKLVD	240	
Qy	241	GSAAKXNSNKNVNA	KLPIGTPDYMABEVLTVMNGDGKGT	YGLBCDWMVSGVIAIYEMTYGR	300	
Db	241	GSAAKXNSNKNVNA	KLPIGTPDYMABEVLTVMNGDGKGT	YGLBCDWMVSGVIAIYEMTYGR	300	
Qy	301	SPAEGTSARTFN	NINMFORFLFPDDPKVSSDFLDLIQ	SLCGQKERLKFEGLCCHPFF	360	
Db	301	SPAEGTSARTFN	NINMFORFLFPDDPKVSSDFLDLIQ	SLCGQKERLKFEGLCCHPFF	360	
Qy	361	SKIDWNINRSP	PPFVPTLKSDDDTNFDPEPKNSWVSS	PCOLSPSGSGEELPVGFS	420	
Db	361	SKIDWNINRSP	PPFVPTLKSDDDTNFDPEPKNSWVSS	PCOLSPSGSGEELPVGFS	420	
Qy	421	YSKALGTGLGR	SESVVSGLDSPAKTSSMEKLLIKSK	ELQDSQDKCHQWQEMTLHRRVS	480	
Db	421	YSKALGTGLGR	SESVVSGLDSPAKTSSMEKLLIKSK	ELQDSQDKCHQWQEMTLHRRVS	480	
Qy	481	EVHVLVSQKEV	ELKASPTORSLLLEQDLATYITECSS	LKRSLEQARMEVSGEDKALQLLH	540	
Db	481	EVHVLVSQKEV	ELKASPTORSLLLEQDLATYITECSS	LKRSLEQARMEVSGEDKALQLLH	540	
Qy	541	DIREQSRKLOE	IKEOEYCAQVEMRLMWNQLEEDLV	SARRSDLYSELSRESLSLAEEPK	600	
Db	541	DIREQSRKLOE	IKEOEYCAQVEMRLMWNQLEEDLV	SARRSDLYSELSRESLSLAEEPK	600	
Qy	601	RKATECHKLLK	KADQCKPVGAYALEKINASOQLKI	OELQSKLKA-----	648	
Db	601	RKATECHKLLK	KADQCKPVGAYALEKINASOQLKI	OELQSKLKA-----	648	
Qy	649	-----	AKERAARELEKLN	RSDSSEGRKKLVEAERRHSLSEN	VKKLETMERENRKLKOD	704
Db	661	NIQAKERAPARE	LEKLNQRSDSSEGRKKLVEAERRHSLSEN	VKKLETMERENRKLKOD	720	
Qy	705	IQTKSQOIQQ	WADKILELEKHXREAOVSAQHLE	VHLKQKEQHYEKKVLJDNQIKKDLAD	764	
Db	721	IQTKSQOIQQ	WADKILELEKHXREAOVSAQHLE	VHLKQKEQHYEKKVLJDNQIKKDLAD	780	
Qy	765	KETLENMMQR	HEBEAHEKGIKLSQKAMINAMDSK	TRLSQRIVELSEANKLAANSLSFT	824	
Db	781	KETLENMMQR	HEBEAHEKGIKLSQKAMINAMDSK	TRLSQRIVELSEANKLAANSLSFT	840	
Qy	825	QRNKAAQEMI	SRLROQKFYLETOAGKLEAQNKL	BEQLEKISHQHSDDXNRILLETRL	884	
Db	841	QRNKAAQEMI	SRLROQKFYLETOAGKLEAQNKL	BEQLEKISHQHSDDXNRILLETRL	900	
Qy	885	REVSLEHEE	QKLEUKROLTELQSLQRESOL	TALQAARAALSQLRQAATELEETAE	944	
Db	901	REVSLEHEE	QKLEUKROLTELQSLQRESOL	TALQAARAALSQLRQAATELEETAE	960	
Qy	945	EEBIIQALT	AHRDEIQRFKDALRNSCTVITOLE	BQNLQJTDENALNNFYLSKQDLEAS	1004	
Db	961	EEBIIQALT	AHRDEIQRFKDALRNSCTVITOLE	BQNLQJTDENALNNFYLSKQDLEAS	1020	

Qy	1005	GANDEIVOLRSEVDHLRREITEREMQLTQKQTMELAKTTCTMLEQVMDLEALNDELLE	1064
Db	1021	GANDEIVOLRSEVDHLRREITEREMQLTQKQTMELAKTTCTMLEQVMDLEALNDELLE	1080
Qy	1065	KERQWEAWSVLGDEKSQFECRVRELQRMJLDTBKQSPARADQIRITSRQVVELAVXHKHA	1124
Db	1081	KERQWEAWSVLGDEKSQFECRVRELQRMJLDTBKQSPARADQIRITSRQVVELAVXHKHA	1140
Qy	1125	EIILALQALKEQKLKAESISDKLNDLEKKHAMLEMNARSLOQKLETERELKORLLSEQAK	1184
Db	1141	EIILALQALKEQKLKAESISDKLNDLEKKHAMLEMNARSLOQKLETERELKORLLSEQAK	1200
Qy	1185	LOQOMDLQKNHIIPLRLTQGLQEALDRADLLKTERSDLEYQLENTQVLYSHEKVMQEGTISQ	1244
Db	1201	LOQOMDLQKNHIIPLRLTQGLQEALDRADLLKTERSDLEYQLENTQVLYSHEKVMQEGTISQ	1260
Qy	1245	QTKLIDFLOAKMDQPAKKKKGLFSRRKEDPALPTQVPLQVNEJLKALEKARKCAELEEA	1304
Db	1261	QTKLIDFLOAKMDQPAKKK-----VPLQVNEJLKALEKARKCAELEEA	1305
Qy	1305	LQKTRIELRSARBEAAHRKATDHPHSTPATARQQTAMSAIVSRPEHOPSAMSLAPPSS	1364
Db	1306	LQKTRIELRSARBEAAHRKATDHPHSTPATARQQTAMSAIVSRPEHOPSAMSLAPPSS	1365
Qy	1365	RRKESSTPEBFSRRLKERVHHNIIPHFPNVGLNWRATKCAVCLDTHVHFGROASKCLEQVM	1424
Db	1366	RRKESSTPEBFSRRLKERVHHNIIPHFPNVGLNWRATKCAVCLDTHVHFGROASKCLEQVM	1425
Qy	1425	CHPKCSTCLPATCGLPAEYATHFEAFCDKXNSPGLQTKEPSSSLHLEGMKVPNNKR	1484
Db	1426	CHPKCSTCLPATCGLPAEYATHFEAFCDKXNSPGLQTKEPSSSLHLEGMKVPNNKR	1485
Qy	1485	GOQGWDRKYIYVLGSKVLIYDNEARAGORPVBEFELCLPDGQVSTHGAVGASELANTAK	1544
Db	1486	GOQGWDRKYIYVLGSKVLIYDNEARAGORPVBEFELCLPDGQVSTHGAVGASELANTAK	1545
Qy	1545	A-----EKAEADAKL	1554
Db	1546	ADVPYILKMWSEPHITCTWPGRTIYLLAPSPDQKQWVTALESVVAGGRVSREKAEADAKL	1605
Qy	1555	LGNSSLKLEGGDLDMNCTLPFSDQVVLVGTESGLYALNVLKNSLTHVPQIGAVFQIYII	1614
Db	1606	LGNSSLKLEGGDLDMNCTLPFSDQVVLVGTESGLYALNVLKNSLTHVPQIGAVFQIYII	1665
Qy	1615	KDLEKLMIAGEBERALCLVDVKVKOSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN	1674
Db	1666	KDLEKLMIAGEBERALCLVDVKVKOSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN	1725
Qy	1675	GLCICAAMPSSKWIILRYNENLSKYCIKRIETSEPCSIHFTNYSILIGNKKEFYEDMKQ	1734
Db	1726	GLCICAAMPSSKWIILRYNENLSKYCIKRIETSEPCSIHFTNYSILIGNKKEFYEDMKQ	1785
Qy	1735	YTLIEEFLDKNDHSLAPAVFAASSNPVSTVOVNSAGOREYLLCFHEFGVFVDSYGRRS	1794
Db	1786	YTLIEEFLDKNDHSLAPAVFAASSNPVSTVOVNSAGOREYLLCFHEFGVFVDSYGRRS	1845
Qy	1795	RTDDLKWSRLPLAFAYREPFLVFTHFNLSLEBIBIQARSSAGTTPARAYLDIPNPRYLGPAL	1854
Db	1846	RTDDLKWSRLPLAFAYREPFLVFTHFNLSLEBIBIQARSSAGTTPARAYLDIPNPRYLGPAL	1905
Qy	1855	SSGAIIYLAASYQDKLRIVCCKGNLVKESGTEHHRGPGSTSSSSNKGCPPTNYNEHITKRVA	1914
Db	1906	SSGAIIYLAASYQDKLRIVCCKGNLVKESGTEHHRGPGSTSSSSNKGCPPTNYNEHITKRVA	1965
Qy	1915	SSPAPPEGSPHREPSTPHRYREGRTLELRDKSPGRPLEREKSPGRMLSTRERSPGRLF	1974
Db	1966	SSPAPPEGSPHREPSTPHRYREGRTLELRDKSPGRPLEREKSPGRMLSTRERSPGRLF	2025
Qy	1975	EDSSRGLPLAGAVRTPLSQVKNRGQSA 2002	
Db	2026	EDSSRGLPLAGAVRTPLSQVKNVMDQSS 2053	

Matches 1981;	Conservative	2;	Mismatches	4;	Indels	81;	Gaps	3;
RESULT 3								
AAO26959								
ID	AAO26959	standard; protein;	2054	AA.				
XX	AC							
XX	AAO26959;							
DT	01-MAY-2003	(first entry)						
XX								
DE	Human CRiK	protein sequence, SEQ ID No 2.						
XX								
KW	Anorectic; hypotensive;	cardiant; antilipaemic; cerebroprotective;						
KW	antigout; osteopathic;	antiarthritic; cytostatic; antidepressant;						
KW	immunomodulator; antimanic;	tranquilliser; antiparkinsonian; nootropic;						
KW	neuroprotective; antinflammatory;	antidiabetic; analgesic;						
KW	human citron rho/rac-interacting kinase;	enzyme; CRiK; ameliorating;						
KW	obesity; comorbidities; cancer;	anorexia; cachexia; bulimia;						
KW	central nervous system disorder;	chronic obstructive pulmonary disease;						
KW	diabetes; pain.							
XX								
OS	Homo sapiens.							
XX								
PN	WO2003004523-A1.							
XX								
PD	16-JAN-2003.							
XX								
PF	28-JUN-2002; 2002WO-BP007156.							
XX								
PR	02-JUL-2001; 2001US-0301841P.							
PR	11-DEC-2001; 2001US-0338651P.							
PR	25-APR-2002; 2002US-0375014P.							
XX								
PA	(PARB) BAYER AG.							
XX								
PI	Zhu Z;							
XX								
DR	WPI; 2003-221576/21.							
XX	N-PSDB; AAL55214.							
XX								
PT	New human citron rho/rac-interacting kinase (CRiK) polypeptide and							
PT	polynucleotide, useful in preventing, ameliorating or treating diseases							
PT	associated with human CRiK dysfunction, e.g. obesity, diabetes or							
PT	Alzheimer's disease.							
XX								
PS	Claim 1; Fig 2; 237pb; English.							
XX								
CC	The invention relates to an isolated polynucleotide encoding a human							
CC	citron rho/rac-interacting kinase polypeptide. The isolated							
CC	polynucleotide comprises a 6165 or 8603 base pair sequence, given in the							
CC	specification. The human citron rho/rac-interacting kinase (CRiK)							
CC	polypeptide and polynucleotide are useful in preventing, ameliorating, or							
CC	treating diseases associated with human CRiK dysfunction such as obesity							
CC	and obesity-associated comorbidities (e.g. hypertension, coronary artery							
CC	disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of							
CC	cancer including endometrial, breast, prostate and colon cancer),							
CC	anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood							
CC	disorders, anxiety disorders, Parkinson's disease or Alzheimer's							
CC	disease), chronic obstructive pulmonary disease, or diabetes. These can							
CC	also be used to treat pain associated with the disorders. The human CRiK							
CC	polypeptide is also useful in diagnostic assays or in genetic testing.							
CC	The expression vector or the reagent is useful in preparing a medicament							
CC	for modulating the activity of a human CRiK in a disease, e.g. obesity, a							
CC	central nervous system disorder, or chronic obstructive pulmonary							
CC	disease. The fusion protein is useful for generating antibodies against a							
CC	CRiK polypeptide and for use in various assay systems. The methods are							
CC	useful in producing and detecting the polynucleotide and polypeptide and							
CC	in screening for agents that modulate the activity of the human CRiK							
CC	polypeptide. This sequence represents the human CRiK protein of the							
CC	invention							
XX								
XX	Sequence 2054 AA;							
XX								
XX	Query Match	95.6%;	Score	10032.5;	DB 6;	Length	2054;	
XX	Best Local Similarity	95.8%;						

Qy 1065 KERQWEAWRSVLGDEKSFQFCRVREIQRMLDTEKQSRARADQRTESRQVVELAVKEHKA 1124
Db 1081 KERQWEAWRSVLGDEKSFQFCRVREIQRMLDTEKQSRARADQRTESRQVVELAVKEHKA 1140
Qy 1125 EIALQALKEQKLKABSLDKLNDLEKHAMLEMNARSLOQKLETERELKOFLEBEQAK 1184
Db 1141 EIALQALKEQKLKABSLDKLNDLEKHAMLEMNARSLOQKLETERELKOFLEBEQAK 1200
Qy 1185 LQQQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENIQVLYSHEKVMQETISQ 1244
Db 1201 LQQQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENIQVLYSHEKVMQETISQ 1260
Qy 1245 QTKLIDFLQAKMOPAKKKGLFSRRKEDPALFTQVPLQYNELKLALEKEKARCABLEEA 1304
Db 1261 QTKLIDFLQAKMOPAKKKK-----VPLQYNELKLALEKEKARCABLEEA 1305
Qy 1305 LQKTRIELRSARBEAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSSMSLLAPSS 1364
Db 1306 LQKTRIELRSARBEAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSSMSLLAPSS 1365
Qy 1365 RRKESSTPEFSRRLKERMHNNIHPHPNVGLNWRATKCAVCLDTVHFGRQAKKLEECQVM 1424
Db 1366 RRKESSTPEFSRRLKERMHNNIHPHPNVGLNWRATKCAVCLDTVHFGRQAKKLEECQVM 1425
Qy 1425 CHPKCSCLPATCGLPAEYATHTEAFCDKMNSSPGLQTKPESSSLHLEGMWVPRNNKR 1484
Db 1426 CHPKCSCLPATCGLPAEYATHTEAFCDKMNSSPGLQTKPESSSLHLEGMWVPRNNKR 1485
Qy 1485 GQQGWRKYIVLBSKVLIVDNEARAGORPVEFELCLPDGDVSHGAVGASELANTAK 1544
Db 1486 GQQGWRKYIVLBSKVLIVDNEARAGORPVEFELCLPDGDVSHGAVGASELANTAK 1545
Qy 1545 A-----EKAERADAKL 1554
Db 1546 ADVPYILKMSHPHTCWPORTULYLAFFDPKQWWTALSVVAGGRVSRKARADAKL 1605
Qy 1555 LGNSLLKLEGGDRDLNCTLPFSDQVVLVGTBEGLYALANVLKNSLTHVPGIGAVFIYII 1614
Db 1606 LGNSLLKLEGGDRDLNCTLPFSDQVVLVGTBEGLYALANVLKNSLTHVPGIGAVFIYII 1665
Qy 1615 KDLEKLIMINGEERALCLVDVKKVQSLAQSHLPAQPDISPNIPEAVKCHLFGACKIBN 1674
Db 1666 KDLEKLIMINGEERALCLVDVKKVQSLAQSHLPAQPDISPNIPEAVKCHLFGACKIBN 1725
Qy 1675 GLCICAMPKSVILRYNENLSKYRKEITETSEPCSCIHTFNYSILIGTNKFYEIDMKQ 1734
Db 1726 GLCICAMPKSVILRYNENLSKYRKEITETSEPCSCIHTFNYSILIGTNKFYEIDMKQ 1785
Qy 1735 YTLLEFLDKNDHSLAPFAASNSFPVSIVQNSAGOREEYLLCPHEFGVFVDSYGRRS 1794
Db 1786 YTLLEFLDKNDHSLAPFAASNSFPVSIVQNSAGOREEYLLCPHEFGVFVDSYGRRS 1845
Qy 1795 RTDGLKWSRLPLAFAYREPVLFTHTNSLEVIETIQARSAGTPARAYLDIPNPRVILGPAT 1854
Db 1846 RTDGLKWSRLPLAFAYREPVLFTHTNSLEVIETIQARSAGTPARAYLDIPNPRVILGPAT 1905
Qy 1855 SSGAIVLASSYQDKLVI CCKGNLVKESGTEHHRGPSTSRSPNKRGPPTYNHEHTKRYA 1914
Db 1906 SSGAIVLASSYQDKLVI CCKGNLVKESGTEHHRGPSTSRSPNKRGPPTYNHEHTKRYA 1965
Qy 1915 SSPAPPEGSHPEPSTPHRYREGRTELARDKSPGPLEREKSPGEMLSSTREERSPGRUF 1974
Db 1966 SSPAPPEGSHPEPSTPHRYREGRTELARDKSPGPLEREKSPGEMLSSTREERSPGRUF 2025
Qy 1975 EDSSRGLPAGAVRTPLSQVKNRGQSA 2002
Db 2026 EDSSRGLPAGAVRTPLSQVKNWQDSS 2053

RESULT 4
ABB81927
ID ABB81927 standard; protein; 2054 AA.

XX ABB81927;
AC 10-OCT-2002 (first entry)
DT Human kinase #1.
DE Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;
KW Citron rho-interacting kinase; Gene therapy; mental disorder; cancer.
OS Homo sapiens.
XX WO200259325-A2.
PN 01-AUG-2002.
PD 20-DEC-2001; 2001WO-US050497.
PF 27-DEC-2000; 2000US-0258335P.
PR (LEXI-) LEXICON GENETICS INC.
PA Yu X, Miranda M, Friddle CJ;
PI WPI; 2002-599796/64.
DR N-PSDB; ABQ78870.
XX Novel polynucleotide encoding human proteins that are structurally
PT similar to animal kinases, useful for drug screening, diagnosis, in gene
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
XX applications.
PS Claim 2; Page 39-43; 50pp; English.
XX The invention relates to a novel human protein that shares structural
CC similarity with animal kinases, including serine-threonine kinases,
CC particularly Citron rho-interacting kinases. The proteins of the
CC invention have nootropic and cytostatic activity. The polynucleotides may
CC have a use in gene therapy. The encoded novel polypeptides are useful for
CC generating antibodies, as reagents in diagnostic assays, for identifying
CC other cellular gene products related to NHP and as reagents in assays for
CC screening for compounds that are useful in the treatment of mental,
CC biological or medical disorders and diseases including cancer. The
CC sequence represents a novel human kinase of the invention
SQ Sequence 2054 AA;
Query Match 95.5%; Score 10022.5; DB 5; Length 2054;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;
Qy 1 MLKFKYGARNPLDAGAAEPIASRASRLNLPFQKPPMTQQQMSPLSREGILDALFVIFE 60
Db 1 MLKFKYGARNPLDAGAAEPIASRASRLNLPFQKPPMTQQQMSPLSREGILDALFVIFE 60
Qy 61 ECSQPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVSRLVCGGHFAVQVREKATG 120
Db 61 ECSQPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVSRLVCGGHFAVQVREKATG 120
Qy 121 DIYAMVKKKALIAQOVVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180
Db 121 DIYAMVKKKALIAQOVVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180
Qy 181 DLLSLNRYEDQDENLIQFYLAELILAVSHVLMGYVHRDIKPENLIVDRTHIKLVDF 240
Db 181 DLLSLNRYEDQDENLIQFYLAELILAVSHVLMGYVHRDIKPENLIVDRTHIKLVDF 240
Qy 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMNGDGKGTGLDCDWMWSVGVIAYEMIYGR 300
Db 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMNGDGKGTGLDCDWMWSVGVIAYEMIYGR 300
Qy 301 SPFAEGTSATFTNNIMNFQRLKPPDDPKVSSQFPLDIQSLLCQKRLKFEGLCCHPFF 360

Db 301 SPFAEGTSARTFNNIMNFRLFKFPDDPKVSDFDLDLIQSLLCGQKRLKFEGLCCHPFF 360
Qy 361 SKTDWNNIRNSPPFPVPTTKSDDDTNPNFBPEKNISWSSSPQLSPGSGSEBELFPVGFS 420
Db 361 SKTDWNNIRNSPPFPVPTTKSDDDTNPNFBPEKNISWSSSPQLSPGSGSEBELFPVGFS 420
Qy 421 YSKALGILGRSESVSGSLDSPAKTSMVEKKLLIKSKELQSDQDKCHKMQEOTRLHRRVS 480
Db 421 YSKALGILGRSESVSGSLDSPAKTSMVEKKLLIKSKELQSDQDKCHKMQEOTRLHRRVS 480
Qy 481 EYAVLSQKEVELKASQTSRSLLEQDLATVITECSSLKRSLEQARNEVSQEDDKALQLLH 540
Db 481 EYAVLSQKEVELKASQTSRSLLEQDLATVITECSSLKRSLEQARNEVSQEDDKALQLLH 540
Qy 541 DIREQSKLQEIKEQYQAVQEMRLMMNQLBEDLVASRRSDLYESELRESRLAABEFK 600
Db 541 DIREQSKLQEIKEQYQAVQEMRLMMNQLBEDLVASRRSDLYESELRESRLAABEFK 600
Qy 601 RKATECOHLLKAKOQKPEVGEYAKLEKINAEQQLKIQELQEKLEKA----- 648
Db 601 RKATECOHLLKAKOQKPEVGEYAKLEKINAEQQLKIQELQEKLEKAVKASTEATELLQ 660
Qy 649 ----AKERAERELEKLNREDSEGIKKLVAEABRRHSLENKVLETFMERENRLKDD 704
Db 661 NIRQAKERAERELEKLNREDSEGIKKLVAEABRRHSLENKVLETFMERENRLKDD 720
Qy 705 IQTKSQIQOMADKILBEKREAOVSQOHLVHLKQEQHYEKKIVLDNOIKKDLAD 764
Db 721 IQTKSQIQOMADKILBEKREAOVSQOHLVHLKQEQHYEKKIVLDNQIKKDLAD 780
Qy 765 KETLENNQORHBEAHEKGIILSEQKAMINAMDSKIRSLERQIVELSEANKLAANSUFT 824
Db 781 KETLENNQORHBEAHEKGIILSEQKAMINAMDSKIRSLERQIVELSEANKLAANSUFT 840
Qy 825 QRMKACEBEMISELROCKFYLETQAGKLEAQNRKLEBEQLEKISHODHSDKNLLELETRL 884
Db 841 QRMKACEBEMISELROCKFYLETQAGKLEAQNRKLEBEQLEKISHODHSDKNLLELETRL 900
Qy 885 REVSLHEHQKLELKRQLTELQSLQERESQLTALQAAFAALESQJRAKTELETTAA 944
Db 901 REVSLHEHQKLELKRQLTELQSLQERESQLTALQAAFAALESQJRAKTELETTAA 960
Qy 945 EBEIQALTAHDEIOIKFPALNSCTVITDLBEQNLQLTEDNAELNNQNFYLSKQLEAS 1004
Db 961 EBEIQALTAHDEIOIKFPALNSCTVITDLBEQNLQLTEDNAELNNQNFYLSKQLEAS 1020
Qy 1005 GANDEIVQLRSEVDHLRRIITEREMOLTSQKQTMELKTTCTMLBEEQVMDLEALNDELLE 1064
Db 1021 GANDEIVQLRSEVDHLRRIITEREMOLTSQKQTMELKTTCTMLBEEQVMDLEALNDELLE 1080
Qy 1065 KERQWEANRSVLGDEKQSECFVRELQRMILDTKQSRAPADQRIITESQVVELAVKEHKA 1124
Db 1081 KERQWEANRSVLGDEKQSECFVRELQRMILDTKQSRAPADQRIITESQVVELAVKEHKA 1140
Qy 1125 EILALOQALKEOKLKAEISLSDKLNLEKXHAMLENMARSLOQKLETERELKORLLEBOAK 1184
Db 1141 EILALOQALKEOKLKAEISLSDKLNLEKXHAMLENMARSLOQKLETERELKORLLEBOAK 1200
Qy 1185 LQQQMDLQKNHIFRLTQGLQEALEDRADLLKTERSDLEYOLENIQVLYSHEKVMEGTISQ 1244
Db 1201 LQQQMDLQKNHIFRLTQGLQEALEDRADLLKTERSDLEYOLENIQVLYSHEKVMEGTISQ 1260
Qy 1245 QTKLIDFLOAKWDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALKEKARCALBEA 1304
Db 1261 QTKLIDFLOAKWDQPAKKK-----VPLQYNELKLALKEKARCALBEA 1305
Qy 1305 LQKTRIELRSAREEAHRKATDHPHSTPATARQIAMSIVRSPHQPSAVSLLAPSS 1364
Db 1306 LQKTRIELRSAREEAHRKATDHPHSTPATARQIAMSIVRSPHQPSAVSLLAPSS 1365
Qy 1365 RKESSTPPEFSRRLKERMHNIPIHFNVLNMRATKCAVCLDTVHFGQASKCLECOVM 1424
Db 1366 RKESSTPPEFSRRLKERMHNIPIHFNVLNMRATKCAVCLDTVHFGQASKCLECOVM 1425

RESULT 5
AAU03501
ID AAU03501 standard; protein; 2053 AA.
XX
AC AAU03501;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #1.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032085.
XX
PR 24-NOV-1999; 99US-0167482P.
XX
PA (SUGE-) SUGEN INC.
XX

Qy 1425 CHPKCSCLPATCGLPAPYATHTEATCRDKNNSFGLQTKPSSSLHLEGMKVPNNKR 1484
Db 1426 CHPKCSCLPATCGLPAPYATHTEAFCRDKNNSFGLQTKPSSSLHLEGMKVPNNKR 1485
Qy 1485 GQQGWDKRYIVLEGSKVLLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANAK 1544
Db 1486 GQQGWDKRYIVLEGSKVLLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANAK 1545
Qy 1545 A-----EKAEBAKL 1554
Db 1546 ADVPYILKMESHPTTCTWPGSTLYLLAPSPDKQWVTALESVVAGGRVSREKAEADAKL 1605
Qy 1555 LGNSLLKLEGGDRDLDMNCTLPFSQDVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1614
Db 1606 LGNSLLKLEGGDRDLDMNCTLPFSQDVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1665
Qy 1615 KDLKELMIAGEERALCLVDVVKYKQSLAQSHLPAQDIPNIFEAVKGCFLFGAGKIEN 1674
Db 1666 KDLKELMIAGEERALCLVDVVKYKQSLAQSHLPAQDIPNIFEAVKGCFLFGAGKIEN 1725
Qy 1675 GLCICCAAMPSKVILRYNENLSKYCIKREIETSEPCSCIHFTNYSILIGTNKFEIDMKQ 1734
Db 1726 GLCICCAAMPSKVILRYNENLSKYCIKREIETSEPCSCIHFTNYSILIGTNKFEIDMKQ 1785
Qy 1735 YTLBEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGORBEYLLCFHEFGVFDVSYGRRS 1794
Db 1786 YTLBEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGORBEYLLCFHEFGVFDVSYGRRS 1845
Qy 1795 RTDILKWSRLPLAPAYREPYLIVTHFNSLEVIEIQARSSAGTTPARAVLIDIPNRYLGPAL 1854
Db 1846 RTDILKWSRLPLAPAYREPYLIVTHFNSLEVIEIQARSSAGTTPARAVLIDIPNRYLGPAL 1905
Qy 1855 SSGAIYGLASSYQDQLRVICCKGNLVKESGTEHHRGPSTSSSPNKRGPPTVNEHITKVA 1914
Db 1906 SSGAIYGLASSYQDQLRVICCKGNLVKESGTEHHRGPSTSSSPNKRGPPTVNEHITKVA 1965
Qy 1915 SSPAPPEGSHPRPSTPHRYREGRETELRRDKSPGRFLEREKSPGRMLSTRERSPORLF 1974
Db 1966 SSPAPPEGSHPRPSTPHRYREGRETELRRDKSPGRFLEREKSPGRILSTRERSPARLF 2025
Qy 1975 EDSRGRPLPAGAVRTPLSQVNVKGRGOSA 2002
Db 2026 EDSRGRPLPAGAVRTPLSQVNVKWDQSS 2053

PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI; 2001-343950/36.
DR N-PSDB; AAS06701.
XX

PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX

PS Claim 7; Fig 2; 433pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX

SQ Sequence 2053 AA;

Query Match 95.4%; Score 10012; DB 4; Length 2053;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 2; Mismatches 5; Indels 82; Gaps 4;
QY 1 MLKFKYGARNPLDAGAAEPTASSASRLNLFQFKPPFMTQQQVSPLSREGILDALFVLF 60
DB 1 MLKFKYGARNPLDAGAAEPTASSASRLNLFQFKPPFMTQQQVSPLSREGILDALFVLF 60
QY 61 ECQOPALMKIKVNSFVRK-YSTTIAELQLOPSAKDFEVRSLVGCCHFAEVQVREKAT 119
DB 61 ECQOPALMKIKVNSFVRK-YSTTIAELQLOPSAKDFEVRSLVGCCHFAEVQVREKAT 120
QY 120 GDIYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQVAFODKNHLYIMBYQPG 179
DB 121 GDIYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQVAFODKNHLYIMBYQPG 180
QY 180 GDIYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQVAFODKNHLYIMBYQPG 239
DB 181 GDIYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQVAFODKNHLYIMBYQPG 240
QY 240 FGSAAKNSNKMVNAKLPICPTDYMAPEVLTVNMGDKGTGYLDCDWSVGVYAYEMVYG 299
DB 241 FGSAAKNSNKMVNAKLPICPTDYMAPEVLTVNMGDKGTGYLDCDWSVGVYAYEMVYG 300
QY 300 RSPFAECTSARTNNIMNFORLKFDPDDPKVSSDFDLTQSLCGQKRLKEGLCCHPF 359
DB 301 RSPFAECTSARTNNIMNFORLKFDPDDPKVSSDFDLTQSLCGQKRLKEGLCCHPF 360
QY 360 FSKIDMNNIRNSPPFPVPTLKSDDDTSNDFEPKNSWSSPCQLSPGSGSEELPFVGF 419
DB 361 FSKIDMNNIRNSPPFPVPTLKSDDDTSNDFEPKNSWSSPCQLSPGSGSEELPFVGF 420
QY 420 SYSKALGILGRSSVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHKEQEMTRLHRRV 479
DB 421 SYSKALGILGRSSVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHKEQEMTRLHRRV 480
QY 480 SEVEAVLSQKEVELKASQETORSLEODLATYITECSSLRSLQARMEVSQEDDKALQLL 539
DB 481 SEVEAVLSQKEVELKASQETORSLEODLATYITECSSLRSLQARMEVSQEDDKALQLL 540
QY 540 HDIREQSRKLOEIQEQYQAVQVEEMLMNQLEEDLVSAARRSDLYSELSRESRLAAEEF 599

DB HDIREQSRKLOEIQEQYQAVQVEEMLMNQLEEDLVSAARRSDLYSELSRESRLAAEEF 600
QY KRKATECHKLLKAKDOGKPEVGEYAKLEKINAEQOLKIQELQEKLEKA----- 648
DB KRKATECHKLLKAKDOGKPEVGEYAKLEKINAEQOLKIQELQEKLEKAVKASTATELL 660
QY -----AKERAERELEKLONREDSSEGIRKKLVBAEERRHSHLENKVKLETWERRNRLKD 703
DB QNIRQAKERAERELEKLONREDSSEGIRKKLVBAEERRHSHLENKVKLETWERRNRLKD 720
QY DIQTSQOIQOMADKILELEKHEKREAVSAQHLEVHLKQKQHYVEEKIKVDNQIKDLA 763
DB DIQTSQOIQOMADKILELEKHEKREAVSAQHLEVHLKQKQHYVEEKIKVDNQIKDLA 780
QY DKETLENMQRHEEFAHEKGIKILSEQAMINAMDSKIRSLERIVELSEANKLAANSLF 823
DB DKETLENMQRHEEFAHEKGIKILSEQAMINAMDSKIRSLERIVELSEANKLAANSLF 840
QY TORNMKAQEEEMISELROOKFYLETQAGKLEAQNKRLEEQLEKISHQDHSQKRLLEETR 883
DB TORNMKAQEEEMISELROOKFYLETQAGKLEAQNKRLEEQLEKISHQDHSQKRLLEETR 900
QY LREVSLEHEEOKLEKQLTELQLSLOERSOLTALQAAALLESOLROKTELEETAE 943
DB LREVSLEHEEOKLEKQLTELQLSLOERSOLTALQAAALLESOLROKTELEETAE 960
QY AEEBIOALTARDEIQKFDALRNSCTVITDLBEQNLQLTEDNAELNNQNFYLSKQLEA 1003
DB AEEBIOALTARDEIQKFDALRNSCTVITDLBEQNLQLTEDNAELNNQNFYLSKQLEA 1020
QY SGANDEIVQURSEVDHLRREITEREMOLTSQKQTEALKTCTMLEQVMDLEALNDELL 1063
DB SGANDEIVQURSEVDHLRREITEREMOLTSQKQTEALKTCTMLEQVMDLEALNDELL 1080
QY EKERQWEANRSLVGDSEKQPECRVRELQRLMDTEKQSRARADQRIETBSRQVVELAVKEHK 1123
DB EKERQWEANRSLVGDSEKQPECRVRELQRLMDTEKQSRARADQRIETBSRQVVELAVKEHK 1140
QY AEIILAQALKEQKLEKAEISDKLNDLEKXHAMLENNARSLQOKLETERELKORLLEEQ 1183
DB AEIILAQALKEQKLEKAEISDKLNDLEKXHAMLENNARSLQOKLETERELKORLLEEQ 1200
QY KLOQOMDLQKHIFRLTQGLQEQALDRADLLKTERSDLEYOLENIQVLYSHEKVMQETIS 1243
DB KLOQOMDLQKHIFRLTQGLQEQALDRADLLKTERSDLEYOLENIQVLYSHEKVMQETIS 1260
QY QOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQVNLKLALEKEKARCABLEE 1303
DB QOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQVNLKLALEKEKARCABLEE 1305
QY ALQKTRIELSARBEAAHRAKATDHPSTPATARQOIANSVAVRSPEHQPSAMSLAPPS 1363
DB ALQKTRIELSARBEAAHRAKATDHPSTPATARQOIANSVAVRSPEHQPSAMSLAPPS 1365
QY SRRKESSTPEFSSRLKERMHNNIPHRFNVGLNMRATKCAVCLDTVHFGRAQSKCLCQV 1423
DB SRRKESSTPEFSSRLKERMHNNIPHRFNVGLNMRATKCAVCLDTVHFGRAQSKCLCQV 1425
QY MCHPKCSTCLPATCGLPAEYATHTFEAFCDKXNSPGLOTKEPSSSLHLGWMKVPNNK 1483
DB MCHPKCSTCLPATCGLPAEYATHTFEAFCDKXNSPGLOTKEPSSSLHLGWMKVPNNK 1485
QY RGQOGWDRKYIVLEGSKVLYIDNEAREAGORPVEEFELCLPDGDSVTHGAVGASELANTA 1543
DB RGQOGWDRKYIVLEGSKVLYIDNEAREAGORPVEEFELCLPDGDSVTHGAVGASELANTA 1545
QY KA-----EKAEADAK 1553
DB KADVPYILKMESHPTTWCWPGRTIYLLAPFPDQKRWVTALSVYAGRVSRKAEADAK 1605
QY LLGSHLLKLEGGDLDNKNCTLPFSDQVVLVGTBEGLVALNVLKNSLTHVPQIGAVFOIYI 1613

Db	1606	LLGNSLLKLEGGDDRLDMNCTLP	PSDQVVLVGTBEGLVALNVLNKSLTHVFGIGAVFOIYI	1666
Qy	1614	IKOLEKLLMIAGEERALCLVDV	KVKVQSLAQSHLPAQFDISPNIFEAVKGCHLFGAGKIE	1673
Db	1666	IKOLEKLLMIAGEERALCLVDV	KVKVQSLAQSHLPAQFDISPNIFEAVKGCHLFGAGKIE	1725
Qy	1674	NGLCICAAAPSKVILIRYNENL	SKYCIKRIKEIETSEPCSCIFHTNYSILIGTNKPYEIDMK	1733
Db	1726	NGLCICAAAPSKVILIRYNENL	SKYCIKRIKEIETSEPCSCIFHTNYSILIGTNKPYEIDMK	1785
Qy	1734	QYTLEEFDLKNDHSLAPAVFAA	SSNSPPVSIQVNSAGQREEVLLCPHEFGVFDVSYGRR	1793
Db	1786	QYTLEEFDLKNDHSLAPAVFAA	SSNSPPVSIQVNSAGQREEVLLCPHEFGVFDVSYGRR	1845
Qy	1794	SRTDLLKWSLPLAFAYREPYL	IFVTHNSLEVIEIQARSSAGTAPARAYLDIPNPRYLGP	1853
Db	1846	SRTDLLKWSLPLAFAYREPYL	IFVTHNSLEVIEIQARSSAGTAPARAYLDIPNPRYLGP	1905
Qy	1854	ISSGAIYLASSYODKLRVIC	CKGNLVKESGTEHHRGPSTGRSSPNKRGPPTYNEHITKRV	1913
Db	1906	ISSGAIYLASSYODKLRVIC	CKGNLVKESGTEHHRGPSTGRSSPNKRGPPTYNEHITKRV	1965
Qy	1914	ASSPAPPEGSPHPRPSTPHRY	EGRTTELRRDXSPGRPLERKSPGRMLSTRRERSPGRL	1973
Db	1966	ASSPAPPEGSPHPRPSTPHRY	EGRTTELRRDXSPGRPLERKSPGRMLSTRRERSPGRL	2025
Qy	1974	FEDSSRGRLPAGAVRTPLSO	VNKGROS 2001	
Db	2026	FEDSSRGRLPAGAVRTPLSO	VNKGROS 2053	
RESULT 6				
ABG78362				
ID	ABG78362	standard; protein; 2066	AA.	
XX	AC	ABG78362;		
XX	DT	15-NOV-2002 (first entry)		
XX	DE	Human protein, homologous to kinases, designated NOV3a.		
XX	KW	Human; NOV; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; diabetes; cell signalling; metabolic pathway; cellular receptor; downstream effector; cancer; gene therapy; hypertension; congenital heart defect; aortic stenosis; obesity; infectious disease; anorexia; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; haemophilia; dyslipidemia; vaccine; haematopoietic disease; scleroderma; fertility; immunogen; idiopathic thrombocytopenic purpura; graft versus host disease; Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease; systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy; stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia; pain; alcoholism; transgenic.		
OS		Homo sapiens.		
XX				
PN	WO200226926-A2.			
XX		04-APR-2002.		
XX		27-SEP-2001; 2001WO-US042336.		
XX				
PR	27-SEP-2000; 2000US-0235631P.			
PR	27-SEP-2000; 2000US-0235633P.			
PR	27-SEP-2000; 2000US-0235808P.			
PR	27-SEP-2000; 2000US-0236064P.			
PR	27-SEP-2000; 2000US-0236065P.			
PR	27-SEP-2000; 2000US-0236066P.			
PR	28-SEP-2000; 2000US-0236135P.			
PR	03-OCT-2000; 2000US-0237434P.			
PR	05-OCT-2000; 2000US-0238321P.			
PR	06-OCT-2000; 2000US-0238396P.			
PR	06-OCT-2000; 2000US-0238399P.			

121 DIYAMKVMKKALLAQEQVSFFPEERNILSRSTSPWIPQLQYAFQDKNHLIYLMVEYQPG 180
181 DLLSLNRYEDQDENLQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVDF 240
181 DLLSLNRYEDQDENLQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVDF 240
241 GSAAKMNSKQVNAKPIGTPDDYMAPEVLTVMNGDGKTYGLDCDWMSGVGIAYEMIYGR 300
241 GSAAKMNSK-VNAKLPIGTPDDYMAPEVLTVMNGDGKTYGLDCDWMSGVGIAYEMIYGR 299
301 SPAEGTSARTFNINWFORFLKFPDDPKVSSDFDLIOGLLGGQKRLKFGELCHPFF 360
300 SPAEGTSARTFNINWFORFLKFPDDPKVSSDFDLIOGLLGGQKRLKFGELCHPFF 359
361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDPEPEKNISWSSPCQLSPSGFSGBELPVGFS 420
360 SKIDWNIRNAPFPVPTLKSDDDTSNFDPEPEKNISWSSPCQLSPSGFSGBELPVGFS 419
421 YSALGILGSESVWSGLDSPAFTSWEKKLITKSKELQSDQKCHKMEQEMTRHRRVS 480
420 YSALGILGSESVWSGLDSPAFTSWEKKLITKSKELQSDQKCHKMEQEMTRHRRVS 479
481 EBAVLISQKEVELKASSETQSLLEQDLATVITECSSLSKRSLEQARMEVSQEDDKALQLH 540
480 EBAVLISQKEVELKASSETQSLLEQDLATVITECSSLSKRSLEQARMEVSQEDDKALQLH 539
541 DIREQSRKLOEIKEOBYQAOVEMRLMQLQEBDLVSARRSDLYSELSRESLAABEPK 600
540 DIREQSRKLOEIKEOBYQAOVEMRLMQLQEBDLVSARRSDLYSELSRESLAABEPK 599
601 RKATECOHLLKAKDQKPGVGYAKLEKINASQOLKIOELOKLEKA----- 648
600 RKATECOHLLKAKDQKPGVGYAKLEKINASQOLKIOELOKLEKAVKASTATELLO 659
649 ----AKERAERELEKQNRDSESEGIKKLVEABERRHSLKVKLETMERRENRLKDD 704
660 NIRAQERAERELEKQNRDSESEGIKKLVEABERRHSLKVKLETMERRENRLKDD 719
705 IQTKSOQIQMAQKILELEKHEBAQVSAQHLVHVKOKEQHYEKKIKVLDNQIKKDLAD 764
720 IQTKSOQIQMAQKILELEKHEBAQVSAQHLVHVKOKEQHYEKKIKVLDNQIKKDLAD 779
765 KETLENMQRHEBEAHEKGIKLEQKAMINAMDSKIRSLFORIVELSEANKLAANSSLT 824
780 KETLENMQRHEBEAHEKGIKLEQKAMINAMDSKIRSLFORIVELSEANKLAANSSLT 839
825 QRMKAQEMISLRQOKFYLEQAGKLAQNRKLEFQLEKISHQDHSQNRILLELETEL 884
840 QRMKAQEMISLRQOKFYLEQAGKLAQNRKLEFQLEKISHQDHSQNRILLELETEL 899
885 REVSLSHEEOKLELXQLTLELOLSQERESQALQAAARAALESQURQAQTELETTAA 944
900 REVSLSHEEOKLELXQLTLELOLSQERESQALQAAARAALESQURQAQTELETTAA 959
945 EEHQAHTHRDEIQKFDALRNSCTVITDLEQLNQLTEDNAELNNQNFYLSKQLEAS 1004
960 EEHQAHTHRDEIQKFDALRNSCTVITDLEQLNQLTEDNAELNNQNFYLSKQLEAS 1019
1005 GANDEIVOLRSEVDHLRREITEREMOLTSOKOTMEALKTTCTMLEBQVMDLEALNDELLE 1064
1020 GANDEIVOLRSEVDHLRREITEREMOLTSOKOTMEALKTTCTMLEBQVMDLEALNDELLE 1079
1065 KERQWEAWRSVLGDEKSFQFCRVREIORMLDTEKQSRADQRITESROVVELAVKEHKA 1124
1080 KERQWEAWRSVLGDEKSFQFCRVREIORMLDTEKQSRADQRITESROVVELAVKEHKA 1139
1125 EILALQALKEOKLKAESLSDKLNLEKKHAMELNARSLOQKLETERELKORLLEEQAK 1184
1140 EILALQALKEOKLKAESLSDKLNLEKKHAMELNARSLOQKLETERELKORLLEEQAK 1199
1185 LOQOMDLQKNHIFRLTQGLQEQALDRADLINTERSDLEYOLENIQVLYSHEKVMGEGTISQ 1244
1200 LOQOMDLQKNHIFRLTQGLQEQALDRADLINTERSDLEYOLENIQVLYSHEKVMGEGTISQ 1259

1245 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVFPLOYNELKALEKSKARCAELEEA 1304
1260 QTKLIDFLQAKMDQPAKKKK-----VPLQYNELKALEKSKARCAELEEA 1304
1305 LQKTRIELRSAREBAARHKATDHPHPSTPATARQOIAWSAIVRSPHEQPSAMVSLAPSS 1364
1305 LQKTRIELRSAREBAARHKATDHPHPSTPATARQOIAWSAIVRSPHEQPSAMVSLAPSS 1364
1365 RRKSSSTPEEFSSRLKERMHNIHPHRENVGLNMRATKCAVCLDTVHFGRQASKCLEQVM 1424
1365 RRKSSSTPEEFSSRLKERMHNIHPHRENVGLNMRATKCAVCLDTVHFGRQASKCLEQVM 1424
1425 CHPKCSTCLPATCGLPAEYATHFTEAFCDKQNSPGLQTEPSSSLHLEGWKMVPRNKR 1484
1425 CHPKCSTCLPATCGLPAEYATHFTEAFCDKQNSPGLQTEPSSSLHLEGWKMVPRNKR 1484
1485 GQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEFEFELCLPDGDVSIHGAVGASELANTAK 1544
1485 GQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEFEFELCLPDGDVSIHGAVGASELANTAK 1544
1545 A-----EKAEADAKL 1554
1545 ADVPYILKMHSHPTTCWPGRTLYLLAPSPDKORWYTALESVVAGGRVSRKAEADAKL 1604
1555 LGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTTEGLYALNVKLSLTHVPDGIQAVFQIYI 1614
1605 LGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTTEGLYALNVKLSLTHVPDGIQAVFQIYI 1664
1615 KDLKLLMIAGEERLALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCFLFCAGKIEN 1674
1665 KDLKLLMIAGEERLALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCFLFCAGKIEN 1724
1675 GLCICAMPKSKVILRYNENLSKYCIKREIETSEPCSCIHTNYSLIGTNKFYEIDMKQ 1734
1725 GLCICAMPKSKVILRYNENLSKYCIKREIETSEPCSCIHTNYSLIGTNKFYEIDMKQ 1784
1735 YTLSEFDDKNDHSLAPAVFAASNSFPVSIQVNSAQOREEYLLCFHFGVFDVSYGRRS 1794
1785 YTLSEFDDKNDHSLAPAVFAASNSFPVSIQVNSAQOREEYLLCFHFGVFDVSYGRRS 1844
1795 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIHQARSSAGTARAYLDIPNRYLGPAL 1854
1845 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIHQARSSAGTARAYLDIPNRYLGPAL 1904
1855 SSGAIIYLASSYQDKLRVICCKGNLVKESGTEHHRGTPSTSRSSPNKRGPTTYNEHITKXVA 1914
1905 SSGAIIYLASSYQDKLRVICCKGNLVKESGTEHHRGTPSTSRSSPNKRGPTTYNEHITKXVA 1964
1915 SSPAPPSPGSHPRPSTPHRYREGRTLRDKSPGRPLEREKSPGRMLSTRRERSRGLF 1974
1965 SSPAPPSPGSHPRPSTPHRYREGRTLRDKSPGRPLEREKSPGRMLSTRRERSRGLF 2024
1975 EDSRGRGLPAGAVETPLSQVNKGRGQASOV 2005
2025 EDSRGRGLPAGAVETPLSQVNKGRGQASOV 2055

RESULT 7

ADA05654

ID ADA05654 standard; protein; 2066 AA.

XX ADA05654;

XX ADA05654;

XX 06-NOV-2003 (first entry)

XX Human NOVlg protein SEQ ID NO:14.

XX human; NOVlg; antidiabetic; anorectic; antibacterial; virucide;

XX immunomodulator; cytosolic; nootropic; neuroprotective;

XX antiparkinsonian; antilipemic; gene therapy; human disease;

XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

Db 660 NIRQAKERARELEKLNQRRDSEGRKVLVEAEERHSLNKKVLEWRENRKDD 719
Qy 705 IQTKSQIQQMAKIDLELEKHEAQAQVLAHLEVHLKQEHYEEKI KVLQDQIKKDLAD 764
Db 720 IQTKSQIQQMAKIDLELEKHEAQAQVLAHLEVHLKQEHYEEKI KVLQDQIKKDLAD 779
Qy 765 KETLENMQRHEBEAHEKGI LSEQKAMINAMDSKIRSLFORIVELSEANKLAANSSLET 824
Db 780 KETLENMQRHEBEAHEKGI LSEQKAMINAMDSKIRSLFORIVELSEANKLAANSSLET 839
Qy 825 QRMKAQEEIMISELRQKQFYLETQAGKLEAQNREKLEEQLEKISHQDHSKNRILLELETRL 884
Db 840 QRMKAQEEIMISELRQKQFYLETQAGKLEAQNREKLEEQLEKISHQDHSKNRILLELETRL 899
Qy 885 REVSLEHEOKLEIKROLTELQLSQBRRESQLTALQARAALLESQLRQKTELETTABA 944
Db 900 REVSLEHEOKLEIKROLTELQLSQBRRESQLTALQARAALLESQLRQKTELETTABA 959
Qy 945 EEBIQALTARHDEIQKFDALRNSCTVITDLEQLNQLTEDNAELNNQNFYLSKOLDEAS 1004
Db 960 EEBIQALTARHDEIQKFDALRNSCTVITDLEQLNQLTEDNAELNNQNFYLSKOLDEAS 1019
Qy 1005 GANDEIVQLRSEVDHURREITEREMQITSQKOTMEALKTTCTMLBEQVMDLBADELLE 1064
Db 1020 GANDEIVQLRSEVDHURREITEREMQITSQKOTMEALKTTCTMLBEQVMDLBADELLE 1079
Qy 1065 KERQWEAARVLDGKSOFCRVRREIQRMIDTEKQSRARADORITESROVVELAVKEHA 1124
Db 1080 KERQWEAARVLDGKSOFCRVRREIQRMIDTEKQSRARADORITESROVVELAVKEHA 1139
Qy 1125 EILALQALKEQKLAAPSLSKDLNLEKHAMLENNARSLOQKLETERELKORLLEEQAK 1184
Db 1140 EILALQALKEQKLAAPSLSKDLNLEKHAMLENNARSLOQKLETERELKORLLEEQAK 1199
Qy 1185 LOQQMDLQKXNIFRLTQGLQEAQDRADLLKTERSDILEYQLENIQVLYSHEKVMGRTISO 1244
Db 1200 LOQQMDLQKXNIFRLTQGLQEAQDRADLLKTERSDILEYQLENIQVLYSHEKVMGRTISO 1259
Qy 1245 QTKLIDFLQAKMDQPAKKKKGLFSREKEDPALPTQVPLQVNLKLALEKARCALEBA 1304
Db 1260 QTKLIDFLQAKMDQPAKKK-----VPLQVNLKLALEKARCALEBA 1304
Qy 1305 LQKTRIELRSAREAAHRKATDHPHSTPATARQQTAMSAIVSRPHQSAMSLAPPSS 1364
Db 1305 LQKTRIELRSAREAAHRKATDHPHSTPATARQQTAMSAIVSRPHQSAMSLAPPSS 1364
Qy 1365 RRKESSTPEEFSRLKERMHNNIPHPFNVLNNRATKCAVCLDTHVFGROASKCLEQVM 1424
Db 1365 RRKESSTPEEFSRLKERMHNNIPHPFNVLNNRATKCAVCLDTHVFGROASKCLEQVM 1424
Qy 1425 CHPKSTCLPATCGLPAEYATHTEAFCDKXNNSPGLQTKEPSSSLHLEGWMMKVPNNKR 1484
Db 1425 CHPKSTCLPATCGLPAEYATHTEAFCDKXNNSPGLQTKEPSSSLHLEGWMMKVPNNKR 1484
Qy 1485 GQGWDRKYIIVLGSKVLIYDNEARAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544
Db 1485 GQGWDRKYIIVLGSKVLIYDNEARAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544
Qy 1545 A-----EKABADAKL 1554
Db 1545 ADVPYILKMBSHPHHTCWPGRTLYLAPSPDPKQRWVTALSVAGGRVREKAEADAKL 1604
Qy 1555 LGNSILKLEGGDRDLNNTLPSDQVVLVGTBEGLYALNVLKNSLTHVPGICAVQIYII 1614
Db 1605 LGNSILKLEGGDRDLNNTLPSDQVVLVGTBEGLYALNVLKNSLTHVPGICAVQIYII 1664
Qy 1615 KDLEKLLMIAGEBRALCLVDVKVKQSLAQSHLPAQFDISPNI FFAVKGCHLFGAGKIEN 1674
Db 1665 KDLEKLLMIAGEBRALCLVDVKVKQSLAQSHLPAQFDISPNI FFAVKGCHLFGAGKIEN 1724
Qy 1675 GLCICAAKPSKVILRYNENLSKYCIKREIETSEPCSCIHTFTNYSILIGTNKFYEIDMKQ 1734

Db 1725 GLCICAAKPSKVILRYNENLSKYCIKREIETSEPCSCIHTFTNYSILIGTNKFYEIDMKQ 1784
Qy 1735 YTLBEFLDKNDHSLAPAVFAAASNSFPVSVIVQNSAGQREBYLLCFHEFGVFDVDSYGRRS 1794
Db 1785 YTLBEFLDKNDHSLAPAVFAAASNSFPVSVIVQNSAGQREBYLLCFHEFGVFDVDSYGRRS 1844
Qy 1795 RTDDLKWSRLPLAFAYREPVLFTVTHFNSLEVIEIQARSSAGTTPARAYLDIPNPRYLGPAL 1854
Db 1845 RTDDLKWSRLPLAFAYREPVLFTVTHFNSLEVIEIQARSSAGTTPARAYLDIPNPRYLGPAL 1904
Qy 1855 SSGAIYLASSYQDKLRLVICCKGNLVKESGTEHHRGPTSSSPNKRGPPTTYNEHITKRV 1914
Db 1905 SSGAIYLASSYQDKLRLVICCKGNLVKESGTEHHRGPTSSSPNKRGPPTTYNEHITKRV 1964
Qy 1915 SSPAPPGSPHPRPSTPHRYREGRTELRDKSPGRPLEREKSPGRMLSTRERSRSGRLF 1974
Db 1965 SSPAPPGSPHPRPSTPHRYREGRTELRDKSPGRPLEREKSPGRMLSTRERSRSGRLF 2024
Qy 1975 EDSRGRIPAGAVRTPLSQVKNKRGQSASQV 2005
Db 2025 EDSRGRIPAGAVRTPLSQVKNKRGQSASQV 2055
RESULT 8
ID ABG78363 standard; protein; 2053 AA.
XX AC ABG78363;
XX DT 15-NOV-2002 (first entry)
XX DE RHO/RAC-interacting citron kinase-like human protein, designated NOV3b.
XX KW Human; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;
KW atherosclerosis; diabetes; cell signalling; metabolic pathway;
KW cellular receptor; downstream effector; cancer; gene therapy;
KW hypertension; congenital heart defect; aortic stenosis; obesity;
KW infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
KW neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
KW haematopoietic disease; scleroderma; fertility; immunogen;
KW idiopathic thrombocytopenic purpura; graft versus host disease;
KW Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
KW systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
KW stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;
KW pain; alcoholism; transgenic.
OS Homo sapiens.
XX WO200226826-A2.
XX PD 04-APR-2002.
XX PF 27-SEP-2001; 2001WO-US042336.
XX PR 27-SEP-2000; 2000US-0235631P.
PR 27-SEP-2000; 2000US-0235633P.
PR 27-SEP-2000; 2000US-0235808P.
PR 27-SEP-2000; 2000US-0236064P.
PR 27-SEP-2000; 2000US-0236065P.
PR 27-SEP-2000; 2000US-0236066P.
PR 28-SEP-2000; 2000US-0236135P.
PR 03-OCT-2000; 2000US-0237434P.
PR 05-OCT-2000; 2000US-0238321P.
PR 06-OCT-2000; 2000US-0238396P.
PR 06-OCT-2000; 2000US-0238399P.
PR 16-MAR-2001; 2001US-0276667P.
PR 31-MAY-2001; 2001US-0294823P.
PR 12-JUL-2001; 2001US-0304868P.
PR 26-SEP-2001; 2001US-00235631.
XX (CURA-) CURAGEN CORP.
XX Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;

PI Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DW;
PT Burgess CE, Radigaru M, Kekuda R, Spytek KA, Leach MD, Shinkets RA;
XX WPI; 2002-499860/53.
DR N-PSDB; ABS63436.
XX Novel isolated NOVX polypeptides and polynucleotides homologous to
PT attractin, plexin, papin-like family of proteins, useful for treating
PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and
PT stroke.
XX Claim 1; Page 44-45; 309pp; English.
XX The invention discloses the isolated human polypeptides, and
CC polynucleotides encoding them, that have been designated NOVX. The
CC polypeptides, polynucleotides and antibodies are useful in treating or
CC preventing a NOVX-associated disorder which is cardiomyopathy, a
CC atherosclerosis and diabetes in a human, where the disorder is related to
CC cell signal processing and metabolic pathway modulation. They can also be
CC used in determining the presence of, or predisposition to, a disease
CC associated with altered levels of the polypeptides and polynucleotides of
CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for
CC identifying an agent that binds to, or that modulates the expression or
CC activity of the polypeptide, for identifying an agent which is cellular
CC receptor or downstream effector, for treating or preventing a NOVX-
CC associated disorder and as a pharmaceutical composition comprising the
CC polypeptide, polynucleotide or the antibody. The polypeptides and
CC polynucleotides are useful in diagnostic applications (e.g. as a marker
CC for cancerous cells or tissue types) where their amounts are assessed, or
CC for the manufacture of a medicament (e.g. gene therapy) for treating or
CC preventing disorders or syndromes such as hypertension, congenital heart
CC defects, aortic stenosis, obesity, infectious diseases, neurodegenerative disorders,
CC Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,
CC haemophilia, dyslipidemias, haematopoietic diseases, scleroderma,
CC fertility, idiopathic thrombocytopenic purpura, graft versus host
CC diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
CC allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar
CC ataxia, pain and alcoholism. They may also be used as immunogens to
CC produce antibodies specific for the invention, and as vaccines.
CC Transgenic cells containing a NOVX expressing construct are useful to
CC produce non-human transgenic animals for studying the function and/or
CC activity of the NOVX proteins and for identifying and/or evaluating
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC expressing construct are useful to produce non-human transgenic animals
CC for studying the function and/or activity of the NOVX proteins and for
CC identifying and/or evaluating modulators of NOVX protein activity. The
CC sequences presented in ABG78359-ABG78371 are the human NOV1-NOV8 proteins
XX Sequence 2053 AA;
XX Query Match 95.4%; Score 10005; DB 5; Length 2053;
XX Best Local Similarity 95.6%; Pred. No. 0;
XX Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;
QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFPGKPPFMTQQQWSPLSREGILDALFVLFE 60
DB 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFPGKPPFMTQQQWSPLSREGILDALFVLFE 60
QY 61 ECSQPALMKIKHVSFNVRKYSTTIAELQPSAKDFVRSLVGCGHFAEVQVREKATG 120
DB 61 ECSQPALMKIKHVSFNVRKYSTTIAELQPSAKDFVRSLVGCGHFAEVQVREKATG 120
QY 121 DIYAMKVMKKALLAQEQVSFEERNILSRSTSPWIPQLOAFODKXHLVMEYVQPG 180
DB 121 DIYAMKVMKKALLAQEQVSFEERNILSRSTSPWIPQLOAFODKXHLVMEYVQPG 180
QY 181 DLLSLNRYEDQDLDENLIQFYLAELLVAVSHVLMGYVHRDKPENILVDRTGHIKLVD 240
DB 181 DLLSLNRYEDQDLDENLIQFYLAELLVAVSHVLMGYVHRDKPENILVDRTGHIKLVD 240
QY 241 GSAAKVNSKQVNAKPIGTPDYMAPEVLTVMGDQKGTGYGLDCDWSVGVYAYEMIYGR 300
DB 241 GSAAKVNSKQVNAKPIGTPDYMAPEVLTVMGDQKGTGYGLDCDWSVGVYAYEMIYGR 300

DB 241 GSAAKVNSK-VNAKLPFGTDPDYMAPEVLTVMGDQKGTGYGLDCDWSVGVYAYEMIYGR 299
QY 301 SPRAEGTSARTENINMNFQFLKFPDDPKVSSDFLDLIQSLCGQKRLKFEGLCCCHPFF 360
DB 300 SPFAEGTSARTENINMNFQFLKFPDDPKVSSDFLDLIQSLCGQKRLKFEGLCCCHPFF 359
QY 361 SKIDWNIRNSPPFPVPTLKSDDTNSFDEPEKNSWVSSPCQLSPSGFSGEELPFVGF 420
DB 360 SKIDWNIRNAPPFPVPTLKSDDTNSFDEPEKNSWVSSPCQLSPSGFSGEELPFVGF 419
QY 421 YSKALGILGRSES VVSGLDSPAKTSMEKLLIKSKELQSDQKCHKEQWEMRLHRVS 480
DB 420 YSKALGILGRSES VVSGLDSPAKTSMEKLLIKSKELQSDQKCHKEQWEMRLHRVS 479
QY 481 EVEAVLSQKEVELKASETQSLLEQDLATYITECSSLKRSLQEQARMEVSDDKALQLH 540
DB 480 EVEAVLSQKEVELKASETQSLLEQDLATYITECSSLKRSLQEQARMEVSDDKALQLH 539
QY 541 DIREQSKLQIBKEQVQAQVEENRMLMNOQLEEDLVARRRSDLYESELRESLIAAEFK 600
DB 540 DIREQSKLQIBKEQVQAQVEENRMLMNOQLEEDLVARRRSDLYESELRESLIAAEFK 599
QY 601 RKATECHQKLLKAKDQKPEVGEVAKLEKINAEQOLKIQELQEKLEK----- 648
DB 600 RKATECHQKLLKAKDQKPEVGEVAKLEKINAEQOLKIQELQEKLEKAVKASTEATLQ 659
QY 649 ----AKERAERELEKLNQNRDSSGIRKKLVEABERHSHLENKVKRLTMMERENRLKDD 704
DB 660 NIRQAKERAERELEKLNQNRDSSGIRKKLVEABERHSHLENKVKRLTMMERENRLKDD 719
QY 705 IQTSQIQIQOMADKILEEKHREAOVSAOHLVHLKQKQHYEETIKVLNDQIKKDLAD 764
DB 720 IQTSQIQIQOMADKILEEKHREAOVSAOHLVHLKQKQHYEETIKVLNDQIKKDLAD 779
QY 765 KETLENMQRHEEAEHKGKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSFLT 824
DB 780 KETLENMQRHEEAEHKGKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSFLT 839
QY 825 QRNKAQBEIMISELRQKQFVLETOAGKLEAQNRKLEBEQLEKISHQDSDKNRLLLETRL 884
DB 840 QRNKAQBEIMISELRQKQFVLETOAGKLEAQNRKLEBEQLEKISHQDSDKNRLLLETRL 899
QY 885 REVLSLEHEQKLEKRLQTLQLSQERESQALQAARAALESQLRQAKTELETTAEA 944
DB 900 REVLSLEHEQKLEKRLQTLQLSQERESQALQAARAALESQLRQAKTELETTAEA 959
QY 945 EEEIQALTARHDEIQRFKDALRNSCTVITDLEQNLQNTEDNAELNNQNFYLSKQLEAS 1004
DB 960 EEEIQALTARHDEIQRFKDALRNSCTVITDLEQNLQNTEDNAELNNQNFYLSKQLEAS 1019
QY 1005 GANDEIVQLRSEVDHLREITEREMQLTSSQOTWEALKTKTCTMLEQVMDLEALNDELLE 1064
DB 1020 GANDEIVQLRSEVDHLREITEREMQLTSSQOTWEALKTKTCTMLEQVMDLEALNDELLE 1079
QY 1065 KERQWEAWRSVLGDEKSFQFCRVRELQRLDTEKQSRARADQRIITESQVVELAVKEHKA 1124
DB 1080 KERQWEAWRSVLGDEKSFQFCRVRELQRLDTEKQSRARADQRIITESQVVELAVKEHKA 1139
QY 1125 EILALQALKEQKLEKSLDKLNDLEKKHMLEMNAARSLQOQLETERELKQRLLEBQAK 1184
DB 1140 EILALQALKEQKLEKSLDKLNDLEKKHMLEMNAARSLQOQLETERELKQRLLEBQAK 1199
QY 1185 LQQQMDLQKNHIFRLTQGLQEQALDRADLLKTERSDLEYQLENIOVLYSHEKVKMEGTISQ 1244
DB 1200 LQQQMDLQKNHIFRLTQGLQEQALDRADLLKTERSDLEYQLENIOVLYSHEKVKMEGTISQ 1259
QY 1245 QTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPOYNELKLALEKQKARCAELEEA 1304
DB 1360 QTKLIDFLQAKMDQPAKKK-----VPLQYNELKLALEKQKARCAELEEA 1304
QY 1305 LQKTRIELRSAREEAAHKKATDHPHPSTPATROQIAMSALVRSPEHQPAMSLLAPSS 1364
DB 1305 LQKTRIELRSAREEAAHKKATDHPHPSTPATROQIAMSALVRSPEHQPAMSLLAPSS 1364

Qy	1365	RRKESSTPEEFSRLKPERMHNIPHRFNVGLNMRATKCAVCLDTHVHFRQASKLECQVM	1424	PF	02-OCT-2002;	2002WO-US031373.
Db	1365	RRKESSTPEEFSRLKPERMHNIPHRFNVGLNMRATKCAVCLDTHVHFRQASKLECQVM	1424	XX	02-OCT-2001;	2001US-0326483P.
Qy	1425	CHPKCSTCLPATCGLPAEYATHTEAPCRDKMNSPGLQTKPESSSLHLEGWKKVPRNNKR	1484	PR	05-OCT-2001;	2001US-0327435P.
Db	1425	CHPKCSTCLPATCGLPAEYATHTEAPCRDKMNSPGLQTKPESSSLHLEGWKKVPRNNKR	1484	PR	05-OCT-2001;	2001US-0327449P.
Qy	1485	GOQGWDRKYIVLEGSKVLIIDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAK	1544	PR	09-OCT-2001;	2001US-0327917P.
Db	1485	GOQGWDRKYIVLEGSKVLIIDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAK	1544	PR	09-OCT-2001;	2001US-0328029P.
Qy	1545	A-----EKAADAKL	1554	PR	09-OCT-2001;	2001US-0328044P.
Db	1545	ADVPYILKMESHPIHTTCWPGRTLYLLAPSPDQKRWVTALESVVAGGRVSREKAADAKL	1604	PR	12-OCT-2001;	2001US-0328056P.
Qy	1555	LGNLSLLKLEGGDRDLNMTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYII	1614	PR	15-OCT-2001;	2001US-0329414P.
Db	1605	LGNLSLLKLEGGDRDLNMTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYII	1664	PR	17-OCT-2001;	2001US-0330142P.
Qy	1615	KDLEKLMIAGERALCLVDVKKVQSLSAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN	1674	PR	18-OCT-2001;	2001US-0330309P.
Db	1665	KDLEKLMIAGERALCLVDVKKVQSLSAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN	1724	PR	22-OCT-2001;	2001US-0341058P.
Qy	1675	GLCICAAAPSKVILRYNENLSKYCIKETESEPCSCITHFTNYSILIGTNKPYEIDMKQ	1734	PR	24-OCT-2001;	2001US-0339266P.
Db	1725	GLCICAAAPSKVILRYNENLSKYCIKETESEPCSCITHFTNYSILIGTNKPYEIDMKQ	1784	PR	24-OCT-2001;	2001US-0343629P.
Qy	1735	YTLLEFLDKNDHSLAPAVFAASSNFPVSIQVNSAGQREYLLCFFHEFGVFDVSYGRS	1794	PR	29-OCT-2001;	2001US-0349575P.
Db	1785	YTLLEFLDKNDHSLAPAVFAASSNFPVSIQVNSAGQREYLLCFFHEFGVFDVSYGRS	1844	PR	01-NOV-2001;	2001US-0346357P.
Qy	1795	RTDDLKWSRLPLAFAYREPFLVFTHTNSLEVBIEIQARSAGTAPARAYLDIPNRYLGPAP	1854	PR	17-APR-2002;	2002US-0373260P.
Db	1845	RTDDLKWSRLPLAFAYREPFLVFTHTNSLEVBIEIQARSAGTAPARAYLDIPNRYLGPAP	1904	PR	19-APR-2002;	2002US-0373815P.
Qy	1855	SSGAIVLASSYQDKLAVICCKGNLVKESGTEHHRGPGSTSRSPNKGPGPTYNEHITKVA	1914	PR	16-MAY-2002;	2002US-0381037P.
Db	1905	SSGAIVLASSYQDKLAVICCKGNLVKESGTEHHRGPGSTSRSPNKGPGPTYNEHITKVA	1964	PR	16-MAY-2002;	2002US-0381042P.
Qy	1915	SSPAPPEGSPHREPSTPHRYREGTELRLRDKSPGRPLREKSPGRMLSTRERSPGRLUF	1974	PR	17-MAY-2002;	2002US-0381642P.
Db	1965	SSPAPPEGSPHREPSTPHRYREGTELRLRDKSPGRPLREKSPGRMLSTRERSPGRLUF	2024	PR	28-MAY-2002;	2002US-0383656P.
Qy	1975	EDSSRGLPAGAVRTPLSQVKNKGOSA	2002	PR	29-MAY-2002;	2002US-0383831P.
Db	2025	EDSSRGLPAGAVRTPLSQVKNKWDQSS	2052	PR	25-JUN-2002;	2002US-0391335P.
RESULT 9				XX	01-OCT-2002;	2002US-00262511.
ADA05642				PA	(CURA-) CURAGEN CORP.	
ID	ADA05642	standard; protein; 2053 AA.		XX	Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;	
AC	ADA05642;			PI	Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;	
XX				PI	Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;	
XX				PI	Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;	
DT	06-NOV-2003 (first entry)			PI	Shinkets RA, Rethenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;	
DE	Human NOV1a protein SEQ ID NO:2.			PI	Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;	
XX	human; NOV1a; antidiabetic; anorectic; antibacterial; virucide;			XX	WPI; 2003-381626/36.	
KW	immunomodulator; cytostatic; nootropic; neuroprotective;			DR	N-PSDB; ADA05641.	
KW	antiparkinsonian; antilipemic; gene therapy; human disease;			XX	New NOVX polypeptides and nucleic acids, useful for diagnosing,	
KW	metabolic disorder; diabetes; obesity; infection; cachexia; cancer;			PT	preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,	
KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;			PT	cancer or dyslipidemia, and in chromosome mapping, tissue typing or	
XX	immune disorder; haematopoietic disorder; dyslipidaemia.			PT	pharmacogenomics.	
XX	Homo sapiens.			PS	Claim 1; Page 99-100; 586pp; English.	
XX	WO2003029424-A2.			XX	The present invention describes NOVX proteins, where X can be 1 to 55	
FN	10-APR-2003.			CC	(e.g. NOV1). Also described: (1) a composition comprising a polypeptide	
XX				CC	described above and a carrier; (2) a kit comprising, in one or more	
XX				CC	containers, the composition described above; (3) an isolated nucleic acid	
XX				CC	molecule which encodes a NOVX protein of the invention; (4) a vector	
XX				CC	comprising the nucleic acid molecule described above; (5) a cell	
XX				CC	comprising the above vector; (6) an antibody that immunospecifically	
XX				CC	binds to the polypeptide described above; (7) methods for determining the	
XX				CC	presence or amount of the above polypeptide or nucleic acid molecule in a	
XX				CC	sample; (8) methods for determining the presence of or predisposition to	
XX				CC	a disease associated with altered levels of expression of the above	
XX				CC	polypeptide or nucleic acid molecule in a first mammalian subject; (9) a	
XX				CC	method of identifying an agent that binds to the polypeptide described	
XX				CC	above; (10) a method for identifying a potential therapeutic agent for	
XX				CC	use in treating a pathology that is related to an aberrant expression or	
XX				CC	aberrant physiological interactions of the polypeptide; (11) a method of	
XX				CC	screening for a modulator of activity or of latency or predisposition to	
XX				CC	a pathology associated with the polypeptide; (12) a method for modulating	
XX				CC	the activity of the polypeptide described above; (13) methods of treating	
XX				CC	or preventing a pathology associated with the above polypeptide in a	
XX				CC	mammal; and (14) a method for producing the above polypeptide. NOVX	

CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
CC and antilipemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.
XX
SQ Sequence 2053 AA;

Query Match 95.4%; Score 10005; DB 6; Length 2053;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;
QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFPGGKPPFMTQQQMSPLSRGILDALFVLF 60
DB 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFPGGKPPFMTQQQMSPLSRGILDALFVLF 60
QY 61 ECSQPALMKI KHVSNFVRKSDTIAELOLQPSAKDFEVRSLVGGCHFAEVQVREKATG 120
DB 61 ECSQPALMKI KHVSNFVRKSDTIAELOLQPSAKDFEVRSLVGGCHFAEVQVREKATG 120
QY 121 DIYAMKMKKALLAOEQVSPFEERNILSRSTSPWIPOLQYAFQDKHLYLMERYQPGG 180
DB 121 DIYAMKMKKALLAOEQVSPFEERNILSRSTSPWIPOLQYAFQDKHLYLMERYQPGG 180
QY 181 DLLSLNRYEDQDENLIFYLAELILAVSHVLMGYVHRDIKPNILVDRTHGHLKLVDF 240
DB 181 DLLSLNRYEDQDENLIFYLAELILAVSHVLMGYVHRDIKPNILVDRTHGHLKLVDF 240
QY 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVNMGDKGTGYGLDCDWSVGVYAYEMIYGR 300
DB 241 GSAAKNSNK - VNAKLPIGTPDYMAPEVLTVNMGDKGTGYGLDCDWSVGVYAYEMIYGR 299
QY 301 SPFAETSAITNNINFORFLKFPDDPKVSSDFDLIQSLICGQKERLKEGLCCHPFF 360
DB 300 SPFAETSAITNNINFORFLKFPDDPKVSSDFDLIQSLICGQKERLKEGLCCHPFF 359
QY 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDPEKNSWSSPQLSPGSPGSEELFVGFS 420
DB 360 SKIDMNNIRNA - PPFVPTLKSDDDTSNFDPEKNSWSSPQLSPGSPGSEELFVGFS 419
QY 421 YSKALGILGRSSVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHKMEQETRLHRRVS 480
DB 420 YSKALGILGRSSVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHKMEQETRLHRRVS 479
QY 481 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSLKESLFOARMEVSOEDDKALQLH 540
DB 480 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSLKESLFOARMEVSOEDDKALQLH 539
QY 541 DIREQSKLQEIKEQYQACVEMRLMMNQLBEDLVSAARRSDLYSELSRSLAAEFK 600
DB 540 DIREQSKLQEIKEQYQACVEMRLMMNQLBEDLVSAARRSDLYSELSRSLAAEFK 599
QY 601 RKATECOHLLKAKOQKPEVEYAKLEKINAEQQLKIOELOKLEKA - - - - - 648
DB 600 RKATECOHLLKAKOQKPEVEYAKLEKINAEQQLKIOELOKLEKAVKASTEATELLQ 659
QY 649 - - - - AKERAERELEKLNREDSEGIKKLVAEARRHSHLENKVKLETMERENRLKDD 704
DB 660 NTRQAKERAERELEKLNREDSEGIKKLVAEARRHSHLENKVKLETMERENRLKDD 719
QY 705 IQTKSQOIQOMADKILEBEKREAOVSAQHLEVLKQEQHYBEKIKVLDNQIKDLAD 764
DB 720 IQTKSQOIQOMADKILEBEKREAOVSAQHLEVLKQEQHYBEKIKVLDNQIKDLAD 779
QY 765 KETLENMQRHBEAEKGIKSLSEQKAMINAMDSKIRSLQRIVELSEANKLAANSLSFT 824

DB 780 KETLENMQRHBEAEKGIKSLSEQKAMINAMDSKIRSLQRIVELSEANKLAANSLSFT 839
QY 825 QRNKAQEMISELRQOKFYLETQAGKLEAQRNKLBEQLEKISHQDSDKNRLLLETRL 884
DB 840 QRNKAQEMISELRQOKFYLETQAGKLEAQRNKLBEQLEKISHQDSDKNRLLLETRL 899
QY 885 REVSLHBEQLEKRLQITELQLSQERESQLTALQAARAALSQLRQAKTELEETTAA 944
DB 900 REVSLHBEQLEKRLQITELQLSQERESQLTALQAARAALSQLRQAKTELEETTAA 959
QY 945 EEEIQALTARDEIQRKFDALRNSCTVITDLEQNLQTEDNABLNQNFVLSKQLEAS 1004
DB 960 EEEIQALTARDEIQRKFDALRNSCTVITDLEQNLQTEDNABLNQNFVLSKQLEAS 1019
QY 1005 GANDEIVOLSEVDHLRREITEREMQLTQSQNTVEALKTKTCTMLEBOVMOLEALNDELLE 1064
DB 1020 GANDEIVOLSEVDHLRREITEREMQLTQSQNTVEALKTKTCTMLEBOVMOLEALNDELLE 1079
QY 1065 KERQWEAWRSVLGDEKSOFCRVRLEQRLMDTEKQSRARADQRTESQVQVVELAVKEHKA 1124
DB 1080 KERQWEAWRSVLGDEKSOFCRVRLEQRLMDTEKQSRARADQRTESQVQVVELAVKEHKA 1139
QY 1125 EILALQALKEQKLKASLSKDLNDEKHAMLEMNARSLOQKLETERELKQRLLEQAK 1184
DB 1140 EILALQALKEQKLKASLSKDLNDEKHAMLEMNARSLOQKLETERELKQRLLEQAK 1199
QY 1185 LQQQMDLQKNHIFRLTQGLQELADRADLAKTERSQLEVOLENQVLYSHEKVKMEGTISQ 1244
DB 1200 LQQQMDLQKNHIFRLTQGLQELADRADLAKTERSQLEVOLENQVLYSHEKVKMEGTISQ 1259
QY 1245 QTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPQVPLQVNLKLALEKQKARCAELEEA 1304
DB 1260 QTKLIDFLQAKMDQPAKKK - - - - - VPLQVNLKLALEKQKARCAELEEA 1304
QY 1305 LQKTRIELRSAREBAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPSAMSLAPPSS 1364
DB 1305 LQKTRIELRSAREBAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPSAMSLAPPSS 1364
QY 1365 RKESSTPEFSRRLKERMHNIIPHRFNVLNMRATCAVCLDTVHFGQASKCLEQVM 1424
DB 1365 RKESSTPEFSRRLKERMHNIIPHRFNVLNMRATCAVCLDTVHFGQASKCLEQVM 1424
QY 1425 CHPKXSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKPESSSLHLEGMKVPNNKR 1484
DB 1425 CHPKXSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKPESSSLHLEGMKVPNNKR 1484
QY 1485 GQQQWDRKYIVLEGSKYLIVDNEAREAGQRPVEEFELCLPDGQVSIHGAVGASLANTAK 1544
DB 1485 GQQQWDRKYIVLEGSKYLIVDNEAREAGQRPVEEFELCLPDGQVSIHGAVGASLANTAK 1544
QY 1545 A - - - - - EKAADAKL 1554
DB 1545 ADVPYILKMESHPHHTCWPGRITLYLLAPSPFDKQWVTALESVVAGGRVSRKAEADAKL 1604
QY 1555 LGNSLLIKLEGDRDLDMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFYI 1614
DB 1605 LGNSLLIKLEGDRDLDMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFYI 1664
QY 1615 KDLKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISFNIFEAVKVGCHLFGAGKIEN 1674
DB 1665 KDLKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISFNIFEAVKVGCHLFGAGKIEN 1724
QY 1675 GLCICAAMPKSVILRYNENLSKYCIKKEITETSEPCSIHFTNYISILIGNKFFEIDMKQ 1734
DB 1725 GLCICAAMPKSVILRYNENLSKYCIKKEITETSEPCSIHFTNYISILIGNKFFEIDMKQ 1784
QY 1735 YTLSEFLDKNDHSLAPAVFAASNSPPVSIVQNSAGOREYLLCFHEFGVFDVSYGRRS 1794
DB 1785 YTLSEFLDKNDHSLAPAVFAASNSPPVSIVQNSAGOREYLLCFHEFGVFDVSYGRRS 1844
QY 1795 RTDLDKNSRLPLAPAYREPFLVTFHNSLEVIETIQARSSAGTAPARAYIDIPNPRYLGPAI 1854

Db 1845 RTDCLKWSRLPLAFAYREPFLVTHFNLSLEVIEIQARSSAGTTPARAYLDIPNPRYLGPAL 1904
 QY 1855 SSGAIIYASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKSGPPTYNEHITKRYA 1914
 Db 1905 SSGAIIYASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKSGPPTYNEHITKRYA 1964
 QY 1915 SSPAPPGSGHPREPSTPHRYREGRTTELKRDKSPGRPLERKSPGRMLSTRERSPGRJLF 1974
 Db 1965 SSPAPPGSGHPREPSTPHRYREGRTTELKRDKSPGRPLERKSPGRMLSTRERSPGRJLF 2024
 QY 1975 EDSRGRPLPAGAVRTPLSQVKNKRGQSA 2002
 Db 2025 EDSRGRPLPAGAVRTPLSQVKNKWDQSS 2052

RESULT 10

ABP97683
 ID ABP97683 standard; protein; 2055 AA.
 XX
 AC ABP97683;
 DT 16-MAY-2003 (first entry)
 XX
 DE Polypeptide similar to citron rho/rac-interacting kinase-short kinase.
 XX
 KW Human; citron rho/rac-interacting kinase-short kinase; obesity;
 KW chronic obstructive pulmonary disease; hypertension; diabetes;
 KW coronary artery disease; hyperlipidaemia; stroke; gallbladder disease;
 KW gout; osteoarthritis; sleep apnea; cancer; thrombotic disease;
 KW polycystic ovarian syndrome; fertility; depression.
 XX

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Best Local Similarity 91.9%; Pred. No. 0;
 Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;
 QY 1 MLKKYKYGARNPLDAGAAEPTASRASRLNLFQGGKPPMTQOQKSPLSREGTLDALFVLEE 60
 Db 1 MLKKYKYGARNPLDAGAAEPTASRASRLNLFQGGKPPMTQOQKSPLSREGTLDALFVLEE 60
 QY 61 ECSOPALMKIKHVNFRVYKYSDDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120
 Db 61 ECSOPALMKIKHVNFRVYKYSDDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120
 QY 121 DIYAMKVMKKALLAQOVSVFFEEERNILSRSTSPWIPQLOYAFODKNHLYLVEEQPGG 180
 Db 121 DIYAMKVMKKALLAQOVSVFFEEERNILSRSTSPWIPQLOYAFODKNHLYLVEEQPGG 180
 QY 181 DLLSLNRYEDQDLENLTQFYLAELIILAVHSHVLMGVVHRDIKPENILVDRTHIKLVDF 240
 Db 181 DLLSLNRYEDQDLENLTQFYLAELIILAVHSHVLMGVVHRDIKPENILVDRTHIKLVDF 240
 QY 241 GSAAKMNSNMVNAKLPIGTDPYNAPEVLTVMGDDGKGTGVLDCDMMWSGVVIAYEYGR 300
 Db 241 GSAAKMNSNMVNAKLPIGTDPYNAPEVLTVMGDDGKGTGVLDCDMMWSGVVIAYEYGR 300
 QY 301 SPRAEGTSARTFNNIMNFQRFKFPDDPKVSSDFDLIQSLLCGQKRLKFEGLCCHPFF 360
 Db 301 SPRAEGTSARTFNNIMNFQRFKFPDDPKVSSDFDLIQSLLCGQKRLKFEGLCCHPFF 360
 QY 361 SKIDWNIRNSPPFPVPTLXSDDDTSNFDPEKNSWVSSPCQLSPSGFSGEELPFVGF 420
 Db 361 SKIDWNIRNSPPFPVPTLXSDDDTSNFDPEKNSWVSSPCQLSPSGFSGEELPFVGF 420
 QY 360 ARTDWNIRNSPPFPVPTLXSDDDTSNFDPEKNSWVSSPCQLSPSGFSGEELPFVGF 419
 Db 360 ARTDWNIRNSPPFPVPTLXSDDDTSNFDPEKNSWVSSPCQLSPSGFSGEELPFVGF 419
 QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSELODQDKCHKEQEMTLHRRVS 480
 Db 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSELODQDKCHKEQEMTLHRRVS 480
 QY 481 EYEAVALSQKEVELKASQTSRSLLEQDLATYITECSSLKRSLQARMEVSQEDDKALQLLH 540
 Db 481 EYEAVALSQKEVELKASQTSRSLLEQDLATYITECSSLKRSLQARMEVSQEDDKALQLLH 540
 QY 541 DIREQSRKLOEIKQOYQAOVEENRMMNQLEEDLVSAARRSDIYSELRESLAAEEFK 600
 Db 541 DIREQSRKLOEIKQOYQAOVEENRMMNQLEEDLVSAARRSDIYSELRESLAAEEFK 600
 QY 540 DIREQSRKLOEIKQOYQAOVEENRMMNQLEEDLVSAARRSDIYSELRESLAAEEFK 599
 Db 540 DIREQSRKLOEIKQOYQAOVEENRMMNQLEEDLVSAARRSDIYSELRESLAAEEFK 599
 QY 601 RKATECOHKLKAKDOQKPEVGEYAKLEKINAEQOLKIQELQKLEKA----- 648
 Db 601 RKATECOHKLKAKDOQKPEVGEYAKLEKINAEQOLKIQELQKLEKA----- 648
 QY 600 RANECOFKLMKAKDOQKPEVGEYAKLEKINAEQOLKIQELQKLEKAVKASTEATELLQ 659
 Db 600 RANECOFKLMKAKDOQKPEVGEYAKLEKINAEQOLKIQELQKLEKAVKASTEATELLQ 659
 QY 649 ----AKERAERELEKLNRRDSSGIRKKLVAEERHSLNKKVRLTETVERRENKDD 704
 Db 649 ----AKERAERELEKLNRRDSSGIRKKLVAEERHSLNKKVRLTETVERRENKDD 704
 QY 660 NIROAKERAERELEKLNRRDSSGIRKKLVAEERHSLNKKVRLTETVERRENKDD 719
 Db 660 NIROAKERAERELEKLNRRDSSGIRKKLVAEERHSLNKKVRLTETVERRENKDD 719
 QY 705 IQTSKQIQQWADKILELEKHEAQAQVSAOHLVHLKQKQHYEKKIKVDNQIKKDLAD 764
 Db 705 IQTSKQIQQWADKILELEKHEAQAQVSAOHLVHLKQKQHYEKKIKVDNQIKKDLAD 764
 QY 720 IQTSKQIQQWADKILELEKHEAQAQVSAOHLVHLKQKQHYEKKIKVDNQIKKDLAD 779
 Db 720 IQTSKQIQQWADKILELEKHEAQAQVSAOHLVHLKQKQHYEKKIKVDNQIKKDLAD 779
 QY 765 KETLENMQRHEEAEHKGKILSEQKAMINAMPSKIRSLRQFIVELSEANKLAANSLSFT 824
 Db 765 KETLENMQRHEEAEHKGKILSEQKAMINAMPSKIRSLRQFIVELSEANKLAANSLSFT 824
 QY 780 KESLENMQRHEEAEHKGKILSEQKAMINAMPSKIRSLRQFIVELSEANKLAANSLSFT 839
 Db 780 KESLENMQRHEEAEHKGKILSEQKAMINAMPSKIRSLRQFIVELSEANKLAANSLSFT 839
 QY 825 QRNMKAQEMISELRQOKFYLETQAGKLEAQNPKLEBQLEKISHQDHSKRNLELETRL 884
 Db 825 QRNMKAQEMISELRQOKFYLETQAGKLEAQNPKLEBQLEKISHQDHSKRNLELETRL 884
 QY 840 QRNMKAQEMISELRQOKFYLETQAGKLEAQNPKLEBQLEKISHQDHSKRNLELETRL 899
 Db 840 QRNMKAQEMISELRQOKFYLETQAGKLEAQNPKLEBQLEKISHQDHSKRNLELETRL 899
 QY 885 REVSLHEEOKLEKRLQTLQSLQERESQTLQAARAALESQURQAATEJETTAA 944
 Db 885 REVSLHEEOKLEKRLQTLQSLQERESQTLQAARAALESQURQAATEJETTAA 944
 QY 900 REVSLHEEOKLEKRLQTLQSLQERESQTLQAARAALESQURQAATEJETTAA 959
 Db 900 REVSLHEEOKLEKRLQTLQSLQERESQTLQAARAALESQURQAATEJETTAA 959
 QY 945 EEEIQALTARDEIQRFDAIRNSCTVITDLEBQLQLTEDNAELANNQNFYLSKQDDEAS 1004
 Db 945 EEEIQALTARDEIQRFDAIRNSCTVITDLEBQLQLTEDNAELANNQNFYLSKQDDEAS 1004
 QY 960 EEEIQALTARDEIQRFDAIRNSCTVITDLEBQLQLTEDNAELANNQNFYLSKQDDEAS 1019
 Db 960 EEEIQALTARDEIQRFDAIRNSCTVITDLEBQLQLTEDNAELANNQNFYLSKQDDEAS 1019
 QY 1005 GANDEIVQLRSEVDHLRREITEREMQVTSQKQTEALKTCTMLEEQVMDLEALNDELLE 1064
 Db 1005 GANDEIVQLRSEVDHLRREITEREMQVTSQKQTEALKTCTMLEEQVMDLEALNDELLE 1064

Db 1020 GANDEIVQLRSEVDHLRRRBITEREMOLTSQKQTMELAKTTCMLBEQVLDLSEALNDELLE 1079
QY 1065 KERQWEAWRSVLGDEKSFQECRVREIQRMLDTEKQSRARADQRIITESQVVELAVKEHKA 1124
Db 1080 KERQWEAWRSVLGDEKSFQECRVREIQRMLDTEKQSRARADQRIITESQVVELAVKEHKA 1139
QY 1125 EIALOQALKEOKLKAESLSDKLNLEKHAMLENNARSLOOKLETERELKORLLEBOAK 1184
Db 1140 EIALOQALKEOKLKAESLSDKLNLEKHAMLENNARSLOOKLETERELKORLLEBOAK 1199
QY 1185 LOQMDLQKNHIFRLTQGLQOALDRADLLKTERSLEYOLENIQVLYSHEKVKMGTTISQ 1244
Db 1200 LOQMDLQKNHIFRLTQGLQOALDRADLLKTERSLEYOLENIQVLYSHEKVKMGTTISQ 1259
QY 1245 QTKLIDFLOAKMDQPAKKKGLFSRKEDPALPTQVPLQYNELKLALEKEKARCABLEEA 1304
Db 1260 QTKLIDFLOAKMDQPAKKK-----VPLQYNELKLALEKEKARCABLEEA 1304
QY 1305 LOKTRIELSAREEAAHRAKATDHPHSTPATARQIQIAMSIVRSPEHOPSANSLILAPSS 1364
Db 1305 LOKTRIELSAREEAAHRAKATDHPHSTPATARQIQIAMSIVRSPEHOPSANSLILAPSS 1364
QY 1365 RKESSTPEFGRRLKERMHNIPIHFNVGLNMRAKCAVCLDTVHFGFQASKCLECQVM 1424
Db 1365 RKESSTPEFGRRLKERMHNIPIHFNVGLNMRAKCAVCLDTVHFGFQASKCLECQVM 1424
QY 1425 CHPKSTCLPATCGIPAEYATHTEAFCDKXNSPGLQTFEPSSSLHLEGMVKVPRNNKR 1484
Db 1425 CHPKSTCLPATCGIPAEYATHTEAFCDKXNSPGLQTFEPSSSLHLEGMVKVPRNNKR 1484
QY 1485 GOQWDRKIVLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544
Db 1485 GOQWDRKIVLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544
QY 1545 A-----EKABADAKL 1554
Db 1545 ADVPYILKMHSHPHITTCWPGRTLYLLAPSPDKQRWVTALESWAGRVSRREKABADAKL 1604
QY 1555 LGNSLLKLGDRDLNMCNLTLPESDQVLVGTTEEGVALNVLKNSLTHVPGIGAVFQIVII 1614
Db 1605 LGNSLLKLGDRDLNMCNLTLPESDQVLVGTTEEGVALNVLKNSLTHVPGIGAVFQIVII 1664
QY 1615 KOLEKLLMIAGBERALCLVDKVKVQSLAQSHLPAQPDIPSPNIFFAVKGCHLFGAGKIEN 1674
Db 1665 KOLEKLLMIAGBERALCLVDKVKVQSLAQSHLPAQPDIPSPNIFFAVKGCHLFGAGKIEN 1724
QY 1675 GLCICAMPKSVVILRYNENLSKYCIKRIETSEPCSCIHFTNYSILIGTGNKFYIDMKQ 1734
Db 1725 SLICICAMPKSVVILRYNENLSKYCIKRIETSEPCSCIHFTNYSILIGTGNKFYIDMKQ 1784
QY 1735 YLLEEFLDKNDHSLAPFAAASNSPPVSVIVQNSAGOREEYLLCFHFEGFVDSYGRRS 1794
Db 1785 YLLEEFLDKNDHSLAPFAAASNSPPVSVIVQNSAGOREEYLLCFHFEGFVDSYGRRS 1844
QY 1795 RTDDLKWSRLPLAFAYREPYLVFVTHNSLEVIETIOARSSAGTPARAYLDIPNRYLGPAT 1854
Db 1845 RTDDLKWSRLPLAFAYREPYLVFVTHNSLEVIETIOARSSAGTPARAYLEIPNRYLGPAT 1904
QY 1855 SSGAIVLASSYQDKLAVICCKGNLVKESGTEHRRGPGSTSRSPKNGPPTTYNEHITKVA 1914
Db 1905 SSGAIVLASSYQDKLAVICCKGNLVKESGTEHRRGPGSTSRSPKNGPPTTYNEHITKVA 1964
QY 1915 SGPAPPEGSPHREPSTPHRY--REGRTLRDKSPGRPLEREKSPGRVLMSTRERSPCR 1972
Db 1965 SGPAPPEGSPHREPSTPHRYDREKRTLRDKSPGRPLEREKSPGRVLMSTRERSPCR 2024
QY 1973 LPEDSRGRPLPAGAVTPLSQVKNKGQSA 2002
Db 2025 LPEDSRGRPLPAGAVTPLSQVKNKWQDQSS 2054

RESULT 11
AA026960

AA026960 standard; protein; 2055 AA.
AA026960;
01-MAY-2003 (first entry)
Human CR1K related protein sequence, SEQ ID No 3.
Anorectic; hypotensive; cardiant; antilipaeamic; cerebroprotective;
antigout; osteopathic; antiarthritic; cytostatic; antidepressant;
immunomodulator; antianic; tranquiliser; antiparkinsonian; nootropic;
neuroprotective; antinflammatory; antidiabetic; analgesic;
human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;
obesity; comorbidities; cancer; anorexia; cachexia; bulimia;
central nervous system disorder; chronic obstructive pulmonary disease;
diabetes; pain.
Homo sapiens.
WO2003004523-A1.
16-JAN-2003.
28-JUN-2002; 2002WO-EP007156.
02-JUL-2001; 2001US-0301841P.
11-DEC-2001; 2001US-0338651P.
25-APR-2002; 2002US-0375014P.
(FARB) BAYER AG.
Zhu Z;
WPI; 2003-221576/21.
New human citron rho/rac-interacting kinase (CR1K) polypeptide and
polynucleotide, useful in preventing, ameliorating or treating diseases
associated with human CR1K dysfunction, e.g. obesity, diabetes or
Alzheimer's disease.
Disclosure; Fig 3; 237pp; English.
The invention relates to an isolated polynucleotide encoding a human
citron rho/rac-interacting kinase polypeptide. The isolated
polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
specification. The human citron rho/rac-interacting kinase (CR1K)
polypeptide and polynucleotide are useful in preventing, ameliorating, or
treating diseases associated with human CR1K dysfunction such as obesity
and obesity-associated comorbidities (e.g. hypertension, coronary artery
disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
cancer including endometrial, breast, prostate and colon cancer),
anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
disorders, anxiety disorders, Parkinson's disease or Alzheimer's
disease), chronic obstructive pulmonary disease, or diabetes. These can
also be used to treat pain associated with the disorders. The human CR1K
polypeptide is also useful in diagnostic assays or in genetic testing.
The expression vector or the reagent is useful in preparing a medicament
for modulating the activity of a human CR1K in a disease, e.g. obesity, a
central nervous system disorder, or chronic obstructive pulmonary
disease. The fusion protein is useful for generating antibodies against a
CR1K polypeptide and for use in various assay systems. The methods are
useful in producing and detecting the polynucleotide and polypeptide and
in screening for agents that modulate the activity of the human CR1K
polypeptide. This sequence represents a protein relating to the human
CR1K protein of the invention
Sequence 2055 AA;

Query Match 92.0%; Score 9656; DB 6; Length 2055;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;
QY 1 MLKFYGARPLDAGAAEPIASRASRLNLPFGQKPPFWTQQKSPLESGILDALFVLE 60

Db 1 MLAFKYGVRNPPASASEPTASASRLNLFQCKPPLMTQOQSALSREGMLDALFALFE 60
Qy 61 ECSQPALMKIKHYSNFKVYSDTIAELQELQPSAKOFVRSVLGCGHFAEVQVVRKATG 120
Db 61 ECSQPALMKIKHYSNFKVYSDTIAELQELQPSAKOFVRSVLGCGHFAEVQVVRKATG 120
Qy 121 DIYAMKVKKALLAQBOVSFFFEERNILSRSTS PMIPOLQYAFQDKNHLXLMEEYOPCG 180
Db 121 DVTAMKIKKALLAQBOVSFFFEERNILSRSTS PMIPOLQYAFQDKNHLXLMEEYOPCG 180
Qy 181 DLSLNLRYEDQDENLQIYLAELILAVHSVHLMGVHRDIKPENILVDRTHIKLVDF 240
Db 181 DFLSLNLRYEDQDENLQIYLAELILAVHSVHLMGVHRDIKPENILVDRTHIKLVDF 240
Qy 241 GSAAKMNSNMVNAKPIGTPDYMAPEVLTVMNGDGKTYGLDCDWMSGVVAYEMVYGR 300
Db 241 GSAAKMNSN- VDAKUPIGTPDYMAPEVLTVMNGDRRGTYGLDCDWMSGVVAYEMVYGR 299
Qy 301 SPFAEGTSARTFNNIMNFQFLKPPDDPKVSSDFLDLIQSLGCGXERLKFEGLCCHPFF 360
Db 300 TPTEGTSARTFNNIMNFQFLKPPDDPKVSSBELLQSLGCGXERLKFEGLCCHPFF 359
Qy 361 SKIDNNIRNSPPFPVPTLKSDDDTSNFDEPEKNSVSSPCOLSPSGSGBELPVGES 420
Db 360 ARTDNNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPFAFSGBELPVGES 419
Qy 421 YSKALGILGRSESVSGSDSPAKTSMEKKLITKSKELQDSQDKHMQEMTRLHRRVS 480
Db 420 YSKALGILGRSESVSGSDSPAKTSMEKKLITKSKELQDSQDKHMQEMTRLHRRVS 479
Qy 481 EVBAVLSQKEVELKASQTSRLLEQDLATYITPCSSLSKRSLEQARMEVSQEDDKALQLLH 540
Db 480 EVBAVLSQKEVELKASQTSRLLEQDLATYITPCSSLSKRSLEQARMEVSQEDDKALQLLH 539
Qy 541 DIREQRKLOEIKBOEYQAOVEBRLMMQLEBDLVSAARRSDLYSELRESRLAAEFK 600
Db 540 DIREQRKLOEIKBOEYQAOVEBRLMMQLEBDLVSAARRSDLYSELRESRLAAEFK 599
Qy 601 RKATECOHKLKAKDQKPEVGYAKLEKINAPQQLKIQELOEKLEKA 648
Db 600 RKANECHKLMAKADQKPEVGYSKLEKINAPQQLKIQELOEKLEKAVKASTATELLIQ 659
Qy 649 ---- AKERAPRELKQNRDSSSEGTRKKLVEABERHSLENKVKLETMERENRLKDD 704
Db 660 NIOKAXERAERELEKLNHRDSSSEGTRKKLVEABERHSLENKVKLETMERENRLKDD 719
Qy 705 IQTKSQIQOMADKILELEEKHREAQVSACHLEVHLKQEQHYEYEEKIKVLDNQIKKDLAD 764
Db 720 IQTKSQIQOMADKILELEEKHREAQVSACHLEVHLKQEQHYEYEEKIKVLDNQIKKDLAD 779
Qy 765 KETLENNQORHEBEAHEKGIKLSQKAMINAMDSKIRSLBOIRIVELSEANKLAANSSUFT 824
Db 780 KESLENNQORHEBEAHEKGIKLSQKAMINAMDSKIRSLBOIRIVELSEANKLAANSSUFT 839
Qy 825 QRNMKAQEMISLRQOKFYLETQAGKLEAQRNKLSEQLEKISHODHSDKNLLELETRL 884
Db 840 QRNMKAQEMISLRQOKFYLETQAGKLEAQRNKLSEQLEKISHODHSDKNLLELETRL 899
Qy 885 REVLSLHEBOKLEKRLKQTELQLSQERESQLTALQARAALLESQLRQAKTELETTAA 944
Db 900 REVLSLHEBOKLEKRLKQTELQLSQERESQLTALQARAALLESQLRQAKTELETTAA 959
Qy 945 EBSIQALTARHDSIQKRFALRNSCTVITDLEBOLNQLTEDNAELNNQNFYLSKOLDEAS 1004
Db 960 EBSIQALTARHDSIQKRFALRNSCTVITDLEBOLNQLTEDNAELNNQNFYLSKOLDEAS 1019
Qy 1005 GANDEIVQLRSEVDHLRREITEREMQLTSSQKQTMELAKTCTCTMLEQVQVMDLBADELLE 1064
Db 1020 GANDEIVQLRSEVDHLRREITEREMQLTSSQKQTMELAKTCTCTMLEQVQVMDLBADELLE 1079
Qy 1065 KERQWEANWSVLGDEKSOFECEVRLEQMLDTEKQSRADQRTITESROQVVELAVKEHKA 1124

Db 1080 KERQWEANWSVLGDEKSOFECEVRLEQMLDTEKQSRADQRTITESROQVVELAVKEHKA 1139
Qy 1125 EILALQALKEOKLKAESLSDKLNDEKKHAMLEMNARSLOQKLETRERLQORLEEOAK 1184
Db 1140 EILALQALKEOKLKAESLSDKLNDEKKHAMLEMNARSLOQKLETRERLQORLEEOAK 1199
Qy 1185 LQQQWMDLQKNHIFRLITQGLQEAALDRADLLKTERSDLEYQLENIQVLYSHKVKWMEGTISQ 1244
Db 1200 LQQQWMDLQKNHIFRLITQGLQEAALDRADLLKTERSDLEYQLENIQVLYSHKVKWMEGTISQ 1259
Qy 1245 QTKLIDFLQAKMDOPAKKXGLFSRRKEDPALPTQVPLQYNELKALEKXKARCAELEEA 1304
Db 1260 QTKLIDFLQAKMDOPAKKXGLFSRRKEDPALPTQVPLQYNELKALEKXKARCAELEEA 1304
Qy 1305 LOKTRIELRSAREEAHRKATDHPHPSTPATARQOIAMSAIVRSPEHOPSAMSLAPPSS 1364
Db 1305 LOKTRIELRSAREEAHRKATDHPHPSTPATARQOIAMSAIVRSPEHOPSAMSLAPPSS 1364
Qy 1365 RRKESSTPEEFSSRLKERMHNIPHRFNVLGNMRAATKCAVCLDTVHFGROASKCLEQVM 1424
Db 1365 RRKESSTPEEFSSRLKERMHNIPHRFNVLGNMRAATKCAVCLDTVHFGROASKCLEQVM 1424
Qy 1425 CHPKCSTCLPATCGLPAEYATHFTEAFCDKMSNPGLOTKEPSSSLHLEGWKMVPRNNKR 1484
Db 1425 CHPKCSTCLPATCGLPAEYATHFTEAFCDKMSNPGLOTKEPSSSLHLEGWKMVPRNNKR 1484
Qy 1485 GQQQWDRKYIVLEGSKVLIYDNEAREAGQRPVSEFELCLPDGDSVHGAVGASELANTAK 1544
Db 1485 GQQQWDRKYIVLEGSKVLIYDNEAREAGQRPVSEFELCLPDGDSVHGAVGASELANTAK 1544
Qy 1545 A-----EKAEADAKL 1554
Db 1545 ADVPYILKMHSHPTTQWPORTLYLLAPSPDKQWTALESVVAGVRGREKAEADAKL 1604
Qy 1555 LGNSLLKLEGDDRLDMNCTLFPDSQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFYII 1614
Db 1605 LGNSLLKLEGDDRLDMNCTLFPDSQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFYII 1664
Qy 1615 KOLEKLMIAAGEERLALCLVDVKVKQSLAOSHUPAOPDI SPNIFEA VKGCHLFGAGKIEN 1674
Db 1665 KOLEKLMIAAGEERLALCLVDVKVKQSLAOSHUPAOPDI SPNIFEA VKGCHLFGAGKIEN 1724
Qy 1675 GLCICAAMPKSVTLRYNENLSKYCIRKETSEPCSCIHFTNYSILIGTNKFYEIDMKQ 1734
Db 1725 SLICICAAMPKSVTLRYNENLSKYCIRKETSEPCSCIHFTNYSILIGTNKFYEIDMKQ 1784
Qy 1735 YTLDEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHFGVFDVSYGRRS 1794
Db 1785 YTLDEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHFGVFDVSYGRRS 1844
Qy 1795 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIQIARSSAGTTPARAYLDTIPNPRYLGPAL 1854
Db 1845 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIQIARSSAGTTPARAYLDTIPNPRYLGPAL 1904
Qy 1855 SSGAIYLAASSYQDKRLVICCKGNLVKESGTHHGRGPTSSRNKNGPPTNYNHIKTRVA 1914
Db 1905 SSGAIYLAASSYQDKRLVICCKGNLVKESGTHHGRGPTSSRNKNGPPTNYNHIKTRVA 1964
Qy 1915 SSPAPPSPGSHPREPSTPHRY--REGTELRDOKSGRPLEREKSPGRMLSTRERSPGR 1972
Db 1965 SSPAPPSPGSHPREPSTPHRY--REGTELRDOKSGRPLEREKSPGRMLSTRERSPGR 2024
Qy 1973 LFEDSSSRGLPAGAVRTPLSQVKNKGRQSSA 2002
Db 2025 LFEDSSSRGLPAGAVRTPLSQVKNKGRQSSA 2054

RESULT 12

ABB81928

ID ABB81928 standard; protein; 1958 AA.

XX

AC ABB81928;

XX

DT 10-OCT-2002 (first entry)
XX Human kinase #2.
XX Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;
KW Citron rho-interacting kinase; gene therapy; mental disorder; cancer.
XX Homo sapiens.
XX WO200259325-A2.
XX 01-AUG-2002.
XX 20-DEC-2001; 2001WO-US050497.
XX 27-DEC-2000; 2000US-02598335P.
XX (LEXI-) LEXICON GENETICS INC.
XX Yu X, Miranda M, Friddle CJ;
XX WPI; 2002-599796/64.
XX N-PSDB; ABQ78871.
XX Novel polynucleotide encoding human proteins that are structurally
PT similar to animal kinases, useful for drug screening, diagnosis, in gene
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
PT applications.
XX Claim 4; Page 46-50; 50pp; English.
XX The invention relates to a novel human protein that shares structural
CC similarity with animal kinases, including serine-threonine kinases,
CC particularly citron rho-interacting kinases. The proteins of the
CC invention have nootropic and cytostatic activity. The polynucleotides may
CC have a use in gene therapy. The encoded novel polypeptides are useful for
CC generating antibodies, as reagents in diagnostic assays, for identifying
CC other cellular gene products related to NHP and as reagents in assays for
CC screening for compounds that are useful in the treatment of mental,
CC biological or medical disorders and diseases including cancer. The
CC sequence represents a novel human kinase of the invention
XX Sequence 1958 AA;
Query Match 90.4%; Score 9497.5; DB 5; Length 1958;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1877; Conservative 1; Mismatches 1; Indels 81; Gaps 3;
QY 1 MLKFKYGARNPLDAGAAEPIASRRLNLFQGGKPPFTQQQMSPLSREGILDALFVLFE 60
DB 1 MLKFKYGARNPLDAGAAEPIASRRLNLFQGGKPPFTQQQMSPLSREGILDALFVLFE 60
QY 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGGCHFAEVQVVRKATG 120
DB 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGGCHFAEVQVVRKATG 120
QY 121 DIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLVLMVEYVPGG 180
DB 121 DIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLVLMVEYVPGG 180
QY 181 DLLSLNRYEDQDENLIFOYLAELILAVHSVHLMGYVHRDIKPNILVDRTGHIKLVD 240
DB 181 DLLSLNRYEDQDENLIFOYLAELILAVHSVHLMGYVHRDIKPNILVDRTGHIKLVD 240
QY 241 GSAAKNSNKNVNAKLPIGTPDYMAPEVLTVNGDGKGTGLDCCDWSVGVIAYEMIYGR 300
DB 241 GSAAKNSNKNVNAKLPIGTPDYMAPEVLTVNGDGKGTGLDCCDWSVGVIAYEMIYGR 300
QY 301 SPFAEGTSARTNNIMNFORFLKFPDDPKVSDFLDLTQSLCCGQKRLKPEGLCCHPFF 360
DB 301 SPFAEGTSARTNNIMNFORFLKFPDDPKVSDFLDLTQSLCCGQKRLKPEGLCCHPFF 360
QY 361 SKIDWNNIRNSPPFPVFTLKSDDDTSNFDPEKNSWSSSPCQLSPGSGSEELPFVGF 420

DB 361 SKIDWNNIRNSPPFPVFTLKSDDDTSNFDPEKNSWSSSPCQLSPGSGSEELPFVGF 420
QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQSDQKCHQVQEMTFLHRV 480
DB 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQSDQKCHQVQEMTFLHRV 480
QY 481 EBEAVLSQKEVELKASRTORSLLSLODLATVITECSSLSKRSLEQARMEVSDDDKALQLH 540
DB 481 EBEAVLSQKEVELKASRTORSLLSLODLATVITECSSLSKRSLEQARMEVSDDDKALQLH 540
QY 541 DIREQSRKLOEIKBOEYQAOVEENRMLMNOLEDLVSAARRSDLYSELSRESRLAAEFK 600
DB 541 DIREQSRKLOEIKBOEYQAOVEENRMLMNOLEDLVSAARRSDLYSELSRESRLAAEFK 600
QY 601 RKATECCHKLLKAKDOCKPGEVAKLEKINAGQKIQELQKLEKXAVKASTATELLQ 660
DB 601 RKATECCHKLLKAKDOCKPGEVAKLEKINAGQKIQELQKLEKXAVKASTATELLQ 660
QY 649 ----AKERAERELEKLNREDSSSEGIKKLVEAEERHSLKVKRLTETMERENRLKDD 704
DB 661 NIRQAKERAERELEKLNREDSSSEGIKKLVEAEERHSLKVKRLTETMERENRLKDD 720
QY 705 IOTXSQIOQWADKILELEBKHREAQVSAQHLEVHLKQKEHYBEKIKVLDNQIKKDLAD 764
DB 721 IOTXSQIOQWADKILELEBKHREAQVSAQHLEVHLKQKEHYBEKIKVLDNQIKKDLAD 780
QY 765 KETLENMQRHEEAEHEKGIKILSEQKAMINAMSKIRSLQRIIVELSEANKLAANSFLT 824
DB 781 KETLENMQRHEEAEHEKGIKILSEQKAMINAMSKIRSLQRIIVELSEANKLAANSFLT 840
QY 825 QRNKAQEEMISELRQOKFYLETOAGKLEAQNRLKEQLEKISHQDSDKNRLLELETRL 884
DB 841 QRNKAQEEMISELRQOKFYLETOAGKLEAQNRLKEQLEKISHQDSDKNRLLELETRL 900
QY 885 REVLSLEHEEOKLEKRLQTELOLSQERESQALQARAALLESOLFQAKTELETTAAE 944
DB 901 REVLSLEHEEOKLEKRLQTELOLSQERESQALQARAALLESOLFQAKTELETTAAE 960
QY 945 EEEIQALTARDEIQRFKDALRNSCTVITDLEQNLQTEEDNALNNQNFYLSKQLEDEAS 1004
DB 961 EEEIQALTARDEIQRFKDALRNSCTVITDLEQNLQTEEDNALNNQNFYLSKQLEDEAS 1020
QY 1005 GANDEIVQLRSEVDHLRREITEREMQLTQKOTMEALKTCTMLEQVMDLEALNDELLE 1064
DB 1021 GANDEIVQLRSEVDHLRREITEREMQLTQKOTMEALKTCTMLEQVMDLEALNDELLE 1080
QY 1065 KERQWEAWRSVLGDSEKQFECRVRELQRLMDTEKQSRARADQRIITBSRQVVELAVKEHKA 1124
DB 1081 KERQWEAWRSVLGDSEKQFECRVRELQRLMDTEKQSRARADQRIITBSRQVVELAVKEHKA 1140
QY 1125 EILALQALKEQKLKABSLSDKLNLEKKGAMLEMMNARSLOQKLETERELKQRLLEBQAK 1184
DB 1141 EILALQALKEQKLKABSLSDKLNLEKKGAMLEMMNARSLOQKLETERELKQRLLEBQAK 1200
QY 1185 LQQQMDLQKHIFRLTOGLOEALDRADLLKATERSDLEYOLENIQVLSYHEKVKMEGTISQ 1244
DB 1201 LQQQMDLQKHIFRLTOGLOEALDRADLLKATERSDLEYOLENIQVLSYHEKVKMEGTISQ 1260
QY 1245 QTKLIDFLQAMPQAKKKGLFSRRKEDPALTPQVPLQYNELKLALKEKARCAELEEA 1304
DB 1261 QTKLIDFLQAMPQAKKKK-----VPLQYNELKLALKEKARCAELEEA 1305
QY 1305 LQKTRIELRSAREEAHRKATDHPSTPATACQIAMSIVBSPHPSAMSLAPPSS 1364
DB 1306 LQKTRIELRSAREEAHRKATDHPSTPATACQIAMSIVBSPHPSAMSLAPPSS 1365
QY 1365 RRKESSTPEEFSSRLKERMEHNIPIHFRNVGLNNRATKCAVCLDTVHFGROASKLECOVM 1424
DB 1366 RRKESSTPEEFSSRLKERMEHNIPIHFRNVGLNNRATKCAVCLDTVHFGROASKLECOVM 1425
QY 1425 CHPKCSTCLPATCGLPABYATHTEAFCDKXKNSPGLQTKPESSSLHLEGMKVPNNKR 1484

Db 1426 CHPKCTCLPATCGLPAEYATHTEAFORDKMNKSPGLQTKPESSSLHLEGWVKVPRNNKR 1485
Qy 1485 GQGWDRKYIVLGRGKVLIDNEAREAGORPVEEFELCLPDGDVSIHGAAGASELANTAK 1544
Db 1486 GQGWDRKYIVLGRGKVLIDNEAREAGORPVEEFELCLPDGDVSIHGAAGASELANTAK 1545
Qy 1545 A-----BKAADAKL 1554
Db 1546 ADVPYILKMBESHPTHCTWPGRTIYLLAPSPDKQRWVTALLESYVAGGRVSRKAEADAKL 1605
Qy 1555 LGSLLKLEGGDRDNNCTLPFSQDVVLVGTBEGLYALVKNLSLTHVPGIGAVFOIYII 1614
Db 1606 LGSLLKLEGGDRDNNCTLPFSQDVVLVGTBEGLYALVKNLSLTHVPGIGAVFOIYII 1665
Qy 1615 KDLEKLMIAAGEBRALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCFLFGAGKIDN 1674
Db 1666 KDLEKLMIAAGEBRALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCFLFGAGKIDN 1725
Qy 1675 GLCICAMPSKVILRYNENLSKYCIKSEIETSEPCSCIHFTNYSILIGTNKPFYIDMKQ 1734
Db 1726 GLCICAMPSKVILRYNENLSKYCIKSEIETSEPCSCIHFTNYSILIGTNKPFYIDMKQ 1785
Qy 1735 YTLLEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHEFGVFDVSYGRRS 1794
Db 1786 YTLLEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHEFGVFDVSYGRRS 1845
Qy 1795 RTDDLKWSRLPLAFAPREPYLFTVTHFNSLEVIQARSSAGTPARAYLDIPNRYLGPAT 1854
Db 1846 RTDDLKWSRLPLAFAPREPYLFTVTHFNSLEVIQARSSAGTPARAYLDIPNRYLGPAT 1905
Qy 1855 SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPFSTGR 1894
Db 1906 SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPFSTGR 1945

RESULT 13
ADD48584
ID ADD48584 standard; protein; 1619 AA.
AC ADD48584;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AF039218, SEQ ID NO 14290.
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
OS
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX GENBANK; AF039218.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)).
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1619 AA;

Query Match 71.7%; Score 7521; DB 7; Length 1619;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1500; Conservative 18; Mismatches 16; Indels 84; Gaps 5;
Qy 468 MEQMTLHRRVSEVAVLSQKEVELKASQTSORSLLEODLATYITCSSLRSLQARME 527
Db 1 MEQEMARLHRRVSEVAVLSQKEVELKASQTSORSLLEODLATYITCSSLRSLQARME 60
Qy 528 VSQEDDKALQLLHDIRQSRKLOEIKQEYQAOVEEMRLMNQLEEDLVSAARRSDIYES 587
Db 61 VSQEDDKALQLLHDIRQSRKLOEIKQEYQAOVEEMRLMNQLEEDLVSAARRSDIYES 120
Qy 588 ELRESRLAAEFKFKATECOHKLKAKDQKPEVGEYAKLEKINAEQOLKIQLQEKLEK 647
Db 121 ELRESRLAAEFKFKATECOHKLKAKDQKPEVGEYAKLEKINAEQOLKIQLQEKLEK 180
Qy 648 A-----AKERAERELEKLNQNRDSSSGIRKLVAEERHSHLENKVKRL 691
Db 181 AVKASTATEALLQNIROAKERAERELEKLNQNRDSSSGIRKLVAEERHSHLENKVKRL 240
Qy 692 ETMERRENRLKDDIQTQSQQIQMADKILELEKHEKRAQVSAQHLEVLHKKQKHVEEKI 751
Db 241 ETMERRENRLKDDIQTQSQQIQMADKILELEKHEKRAQVSAQHLEVLHKKQKHVEEKI 300
Qy 752 KYLDNQIKKDLADKETLENMQRHEEAHEKGLTSPKAMINAMDSKIRSLQRIVELS 811
Db 301 KYLDNQIKKDLADKETLENMQRHEEAHEKGLTSPKAMINAMDSKIRSLQRIVELS 360
Qy 812 EANKLAANSLSFTQNMKAQEFEMISELRQOKFYLETQAGKLEAQNRLKLEPQLEKISHQDH 871
Db 361 EANKLAANSLSFTQNMKAQEFEMISELRQOKFYLETQAGKLEAQNRLKLEPQLEKISHQDH 420
Qy 872 SDKRILLETRELVSEVLEHEEQKLEKROITELQLSLOERESQLTALQARALESQLR 931
Db 421 SDKRILLETRELVSEVLEHEEQKLEKROITELQLSLOERESQLTALQARALESQLR 480
Qy 932 QAKTELEBTTAAEABEEIQAHTAHREDEIQRKFDALRNSCTVITDLEEQNLQTDNAELNN 991
Db 481 QAKTELEBTTAAEABEEIQAHTAHREDEIQRKFDALRNSCTVITDLEEQNLQTDNAELNN 540
Qy 992 QNFYLSKQLEASGANDIEIVQLRSEVDHLRREITEREMQLTQSKQOTWEALKTTCTMLEEQ 1051
Db 541 QNFYLSKQLEASGANDIEIVQLRSEVDHLRREITEREMQLTQSKQOTWEALKTTCTMLEEQ 600

QY 528 VSQEDKALQLLHDIREQSRKLEIQEYQAOVEEMRLMNQLEEDLVASRRSDLYES 587
DB 61 VSQEDKALQLLHDIREQSRKLEIQEYQAOVEEMRLMNQLEEDLVASRRSDLYES 120
QY 588 ELRESLAAEFKFKATECHOKLLKAKDOGKPEVGYAKLEKINABOOLKIOLEQKLEK 647
DB 121 ELRESLAAEFKFKANECHOKLMKAKDLGKPEVGCSEKLEKINABOOLKIOLEQKLEK 180
QY 648 A-----AKERARELEKLEQNREDSSEGIRKXKLVABEERRHLENKVXRL 691
DB 181 AVFASATEATELLQNIQAKERARELEKLEHNREDSSEGIRKXKLVABEERRHLENKVXRL 240
QY 692 ETWERENRLKODIQTKSQIQOMADKILLEBKREAOVSACHLEVLHUKOKEQHYEKEI 751
DB 241 ETWERENRLKODIQTKSQIQOMADKILLEBKREAOVSACHLEVLHUKOKEQHYEKEI 300
QY 752 KVLNDQIKDLADKETLENMORHEEBAHEKGKILSEOKAMINAMDSKIRSLEQRIVELS 811
DB 301 KVLNDQIKDLADKESLETMMQHEEBAHEKGKILSEOKAMINAMDSKIRSLEQRIVELS 360
QY 812 EANKLAANSLFTQNNKQAEEMISELROQKFYLETOAGKLEAQNKLEBEQLEKISHQDH 871
DB 361 EANKLAANSLFTQNNKQAEEMISELROQKFYLETOAGKLEAQNKLEBEQLEKISHQDH 420
QY 872 SDKNRLELETRUREVLSHEEOKLEKQLELOLSQERSOLTAQAAALLESQUR 931
DB 421 SDKNRLELETRUREVLSHEEOKLEKQLELOLSQERSOLTAQAAALLESQUR 480
QY 932 QAKTELEETTAEEBIEIQAHTARDEIQKFDALNSCTVITDLEBQLNQLTDENAELNN 991
DB 481 QAKTELEETTAEEBIEIQAHTARDEIQKFDALNSCTVITDLEBQLNQLTDENAELNN 540
QY 992 QNPLYSKQLDEASGANDEIVQLSEVDHLAREITEREMQLTTSQKQTMELAKTCTMLBEQ 1051
DB 541 QNPLYSKQLDEASGANDEIVQLSEVDHLAREITEREMQLTTSQKQTMELAKTCTMLBEQ 600
QY 1052 VMDLEALNDELLEKQEWAWRSVLGDEKSQFCRVRELQRMLDTEKQSRARADQRIYES 1111
DB 601 VMDLEALNDELLEKQEWAWRSVLGDEKSQFCRVRELQRMLDTEKQSRARADQRIYES 660
QY 1112 QVVELAVKHEKAEIILAOALKEQKKAESLSDKUNLDEKHAMLEMMARSLOQKLETE 1171
DB 661 QVVELAVKHEKAEIILAOALKEQKKAESLSDKUNLDEKHAMLEMMARSLOQKLETE 720
QY 1172 RELKORLEBQAKLOQOMDLQKNHIFRLTQGLQEALDRADLLKTERSDEYOLENIQVLY 1231
DB 721 RELKORLEBQAKLOQOMDLQKNHIFRLTQGLQEALDRADLLKTERSDEYOLENIQVLY 780
QY 1232 SHEKVMKEGTISQOTKLIDFLQAKMDQPAKKKGLFSRKEDPALPTQVFLQYNELKAL 1291
DB 781 SHEKVMKEGTISQOTKLIDFLQAKMDQPAKKKGLFSRKEDPALPTQVFLQYNELKAL 840
QY 1292 EKEKARCAELEALOKTRIELRSAREEAAHRKATDHPHSTPATARQQIAMSALVRSPEH 1351
DB 841 EKEKARCAELEALOKTRIELRSAREEAAHRKATDHPHSTPATARQQIAMSALVRSPEH 900
QY 1352 QPSAMSLAPPSSRRKESSTPEEFSRRLKERMHNNIPHRFNVGLNNRATKCAVCLDTVHF 1411
DB 901 QPSAMSLAPPSSRRKESSTPEEFSRRLKERMHNNIPHRFNVGLNNRATKCAVCLDTVHF 960
QY 1412 GROASKLEQVWCHPKCSTCLPATCGLPAEYATHEAFCDKXNSPGLQTKPESSSLH 1471
DB 961 GROASKLEQVWCHPKCSTCLPATCGLPAEYATHEAFCDKXNSPGLQTKPESSSLH 1020
QY 1472 LEGWMKVPRNNKGGQGWDRKYIVLGSKVLYVDNEAREAGQRPVEFELCLPDGDSVSH 1531
DB 1021 LEGWMKVPRNNKGGQGWDRKYIVLGSKVLYVDNEAREAGQRPVEFELCLPDGDSVSH 1080
QY 1532 GAVGASELANAKA----- 1545
DB 1081 GAVGASELANAKADVPYILKMBESHPHTTCWFGRITLYLLAPSPFDKQWVTALESVWAG 1140
QY 1546 -----EKAEDA-----KLLGNSLLKLEGGDRDLDNMNCTLPFSQVVLVGE 1586

DB 1141 RVSEKAEADAAMDCTCERLPVWVEKLLGNSLLKLEGGDRDLDNMNCTLPFSQVVLVGE 1200
QY 1587 EGLYALANVLKNSLTHVPGIGAVFOIYIIKOLEKLLMIAGEERALCLVDVKKVQSLAQSH 1646
DB 1201 EGLYALANVLKNSLTHVPGIGAVFOIYIIKOLEKLLMIAGEERALCLVDVKKVQSLAQSH 1260
QY 1647 LPAQPDISPNIPEAVKGCHLFGAGKLENGLCICAAMPSKVILRYNENLSKYCIRKEIET 1706
DB 1261 LPAQPDISPNIPEAVKGCHLFGAGKLENGLCICAAMPSKVILRYNENLSKYCIRKEIET 1320
QY 1707 SEPCSCIHFTNYSILITGNKFYEIDMKQYTLERFLDKNDHSLAPAFVAAASNSFPYSIVQ 1766
DB 1321 SEPCSCIHFTNYSILITGNKFYEIDMKQYTLERFLDKNDHSLAPAFVAAASNSFPYSIVQ 1380
QY 1767 VNSAQREEVLLCFHEFGVFDVSYGRSRTDDDKWSRLPLAFAYREPYLVTFHNSLEVI 1826
DB 1381 ANSGQREEVLLCFHEFGVFDVSYGRSRTDDDKWSRLPLAFAYREPYLVTFHNSLEVI 1440
QY 1827 EIQARSSAGTTPARAYLDIPNRYLGPATSSGAYLASSYQDKLRVICCKGNLVKESGTEH 1886
DB 1441 EIQARSSAGTTPARAYLDIPNRYLGPATSSGAYLASSYQDKLRVICCKGNLVKESGTEH 1500
QY 1887 HRGFTSRSSPNKRGPPTYNEHITKRVASSPAPPEGSPHSPREPSTPHRYDRDREGTELRR 1944
DB 1501 HRVFTSR-SPNKGPPTYNEHITKRVASSPAPPEGSPHSPREPSTPHRYDRDREGTELRR 1559
QY 1945 DXSPGRPLEKSPGRMLSTRERSPGRLFEDSSRGLPAGAVRTPLSQVKNKGQOSA 2002
DB 1560 DXSPGRPLEKSPGRMLSTRERSPGRLFEDSSRGLPAGAVRTPLSQVKNKGQOSA 1617

RESULT 15
ID AAB43359 standard; protein; 1286 AA.
XX AAB43359;
DT 08-FEB-2001 (first entry)
XX Human ORFX ORF3123 polypeptide sequence SEQ ID NO:6246.
DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX Homo sapiens.
XX WO200058473-A2.
XX 05-OCT-2000.
XX 31-MAR-2000; 2000WO-US008621.
XX 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.
DR N-PSDB; AAC77568.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 11; Page 5433-5436; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in ABA0237 to ABA43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antiparasitic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotrophic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
XX Sequence 1286 AA;
Query Match 59.3%; Score 6223.5; DB 3; Length 1286;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1231; Conservative 1; Mismatches 3; Indels 65; Gaps 2;
QY 753 VLDNOIKDLADKETLENMOMQHEEAEKGIKSEQKAMINAMDSKIRSEQRIVELSE 812
Db 1 VLDNQIKDLADKETLENMOMQHEEAEKGIKSEQKAMINAMDSKIRSEQRIVELSE 60
QY 813 ANKLAANSLSFTQRMKAQEMISELRQKFKYLETQAGKLEAQRKLEEQLEKISHQDHS 872
Db 61 ANKLAANSLSFTQRMKAQEMISELRQKFKYLETQAGKLEAQRKLEEQLEKISHQDHS 120
QY 873 DKNRLELETRLEVSLEHEEOKLEKQLTELQLSLOERESQLTALQAAALESQLRQ 932
Db 121 DKNRLELETRLEVSLEHEEOKLEKQLTELQLSLOERESQLTALQAAALESQLRQ 180
QY 933 AKTELEETTAABEEIQALTARHDEIQKFDALRNSCTVITDLEEQNLQDTEADNAELNQ 992
Db 181 AKTELEETTAABEEIQALTARHDEIQKFDALRNSCTVITDLEEQNLQDTEADNAELNQ 240
QY 993 NPYLSQKLDASGANDIEIVQLRSEVDHLRREITEREMOLTSQKQTMELAKTCTMLEBEQ 1052
Db 241 NPYLSQKLDASGANDIEIVQLRSEVDHLRREITEREMOLTSQKQTMELAKTCTMLEBEQ 300
QY 1053 MDLEALNDELLEKQEWAEWSVLGDEKSOECREVRLEQLRMLDTEKQSRADQRIETSR 1112
Db 301 MDLEALNDELLEKQEWAEWSVLGDEKSOECREVRLEQLRMLDTEKQSRADQRIETSR 360
QY 1113 QVVELAVKHEKAEIILALQALKEQKKAESLSKLNLEKHEKHAMLENNARSLOQKLETER 1172
Db 361 QVVELAVKHEKAEIILALQALKEQKKAESLSKLNLEKHEKHAMLENNARSLOQKLETER 420
QY 1173 ELKQRLLEBEQAKLOQMDLQKNHIFRLTQGLQOALDRADLLKTERSDEYLENIQVLYS 1232
Db 421 ELKQRLLEBEQAKLOQMDLQKNHIFRLTQGLQOALDRADLLKTERSDEYLENIQVLYS 480
QY 1233 HEKVMWEGTISQOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALE 1292
Db 481 HEKVMWEGTISQOTKLIDFLQAKMDQPAKKK-----VPLQYNELKLALE 525

QY 1293 KEKARCAELEBEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQOIAISAIVRSEHQ 1352
Db 526 KEKARCAELEBEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQOIAISAIVRSEHQ 585
QY 1353 PSAMSLIAPSSRRKESSTPEFSRRLKERMHNHIFHRFVGLNMRATKCAVCLDTVHPG 1412
Db 586 PSAMSLIAPSSRRKESSTPEFSRRLKERMHNHIFHRFVGLNMRATKCAVCLDTVHPG 645
QY 1413 QASKCLCEQVMCHPKCSTCLPATCGI PAEVATHTFAFCRDKVNSPGLQTKPSSSLHL 1472
Db 646 QASKCLCEQVMCHPKCSTCLPATCGI PAEVATHTFAFCRDKVNSPGLQTKPSSSLHL 705
QY 1473 EGMKVPNNKRGQGWDRKYIYLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDSVSHG 1532
Db 706 EGMKVPNNKRGQGWDRKYIYLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDSVSHG 765
QY 1533 AVGASELANATA----- 1545
Db 766 AVGASELANATAKADVPVILKQWESHPHTTWPGBRTLYLLAPSPFPKQWVTALESVWAGR 825
QY 1546 ---EKAADAKLLGNSLLKLEGGDRDLMNCTLPFSQVVLVGTGEGLYALNVLKNSLTHV 1602
Db 826 VSREKADAKLLGNSLLKLEGGDRDLMNCTLPFSQVVLVGTGEGLYALNVLKNSLTHV 885
QY 1603 PGIGAVPQIYIIXDLEKLMIAEERALCLVDVVKVQSLAQSHLPAQDISNIFEAVK 1662
Db 886 PGIGAVPQIYIIXDLEKLMIAEERALCLVDVVKVQSLAQSHLPAQDISNIFEAVK 945
QY 1663 GCHLFGAGKTEGLICICAMPKSVILRYNENLSKYCIKEIETSPSCIHFTNYSILI 1722
Db 946 GCHLFGAGKTEGLICICAMPKSVILRYNENLSKYCIKEIETSPSCIHFTNYSILI 1005
QY 1723 GTNKFYEIDMKQYTLBEFFLDKNDHSLAPAVFAASSNSFPVSIYQVNSAGQREYLLCFHE 1782
Db 1006 GTNKFYEIDMKQYTLBEFFLDKNDHSLAPAVFAASSNSFPVSIYQVNSAGQREYLLCFHE 1065
QY 1783 FGVFVDSYGRSRTDCLKWSRLPLAFAYREPYLFTVTHNSLEVIETQARSAGTAPAYL 1842
Db 1066 FGVFVDSYGRSRTDCLKWSRLPLAFAYREPYLFTVTHNSLEVIETQARSAGTAPAYL 1125
QY 1843 DIPNRYLGPASSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPFSTRSSPNKRG 1902
Db 1126 DIPNRYLGPASSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPFSTRSSPNKRG 1185
QY 1903 PTYNEHITKEVASSPAPPPEGSPHPRPSTPHRYREGRTELRRDKSPGRPLEREKSPGRML 1962
Db 1186 PTYNEHITKEVASSPAPPPEGSPHPRPSTPHRYREGRTELRRDKSPGRPLEREKSPGRML 1245
QY 1963 STRRERSPGRLFEDSSGRPLPAGAVRTPISQVKNKRGQSA 2002
Db 1246 STRRERSPGRLFEDSSGRPLPAGAVRTPISQVKNKWDQSS 1285

Search completed: July 3, 2004, 10:14:52
Job time : 111 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 18:47:57 ; Search time 24219 Seconds
(without alignments)
11765.024 Million cell updates/sec

Title: US-10-017-216-1
Perfect score: 6574
Sequence: 1 agagccgcagtgaggagat.....atcgagaatgtagggttaga 6574

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	6574	100.0	6574	6	AX429512	AX429512 Sequence
2	6159	93.7	6159	6	AX429514	AX429514 Sequence
3	5684.4	86.5	6298	6	AX504254	AX504254 Sequence
4	5668	86.2	8603	6	AX671108	AX671108 Sequence
5	5666.6	86.2	6156	6	AX671112	AX671112 Sequence
6	5666.2	86.2	6165	6	AX671105	AX671105 Sequence
7	5661.4	86.1	6165	6	AX574425	AX574425 Sequence
8	5650.4	86.0	6159	6	AX166510	AX166510 Sequence
9	5631	85.7	6189	6	AX503780	AX503780 Sequence
10	5629.8	85.6	6201	6	AX503778	AX503778 Sequence
11	5552.8	84.5	8576	9	AY257469	AY257469 Homo sapi
12	5358	81.5	5877	6	AX574427	AX574427 Sequence
13	4632.4	70.5	6954	10	AF086824	AF086824 Mus muscu
14	3602.6	54.8	5952	10	AF039218	AF039218 Rattus no
15	3593.8	54.7	4967	10	AF070066	AF070066 Mus muscu
16	3546.4	53.9	5019	10	MM039904	U39904 Mus musculu
17	2436.8	37.1	5261	6	AX671110	AX671110 Sequence
18	2436.8	37.1	5261	9	AB023166	AB023166 Homo sapi
19	2187.4	33.3	3495	10	BC051165	BC051165 Mus muscu
20	2187.4	33.3	3506	10	BC023775	BC023775 Mus muscu
21	1406.4	21.4	2066	6	AX642956	AX642956 Sequence
22	1397.8	21.3	1515	6	AX253937	AX253937 Sequence
23	1395	21.2	1485	6	AX671037	AX671037 Sequence
24	1395	21.2	1765	6	AX671044	AX671044 Sequence
25	1087.8	16.5	2380	6	AX671042	AX671042 Sequence
26	1087.8	16.5	2380	10	AF086823	AF086823 Mus muscu
27	1083.4	16.5	1386	10	AF070065	AF070065 Rattus no
28	768	11.7	1133	6	AB411948	AB411948 Sequence
29	648	9.9	1058	6	AX053315	AX053315 Sequence
30	592	9.0	817	6	AX053416	AX053416 Sequence
31	570	8.7	2390	9	AF022116	AF022116 Homo sapi
32	570	8.7	2395	9	BC017671	BC017671 Homo sapi
33	570	8.7	2402	9	BC001007	BC001007 Homo sapi
34	570	8.7	2429	9	BC001056	BC001056 Homo sapi
35	570	8.7	2429	9	BC001823	BC001823 Homo sapi
36	570	8.7	2442	9	AK127820	AK127820 Homo sapi
37	570	8.7	136436	9	AC002563	AC002563 Human PAC
38	570	8.7	190514	2	AC026363	AC026363 Homo sapi
39	570	8.7	208686	9	AC079262	AC079262 Homo sapi
40	568.4	8.6	4122	9	HS0805803	BX537486 Homo sapi
41	544.8	8.3	2089	9	HS049636	U87276 Human AMP-a
42	544.8	8.3	2353	9	HS039994	U83994 Human AMP-a
43	434	6.6	2284	9	AK123136	AK123136 Homo sapi
44	373.2	5.7	398	6	AX913973	AX913973 Sequence
45	373.2	5.7	398	6	BD049506	BD049506 Sequence

ALIGNMENTS

RESULT 1
AX429512
LOCUS AX429512 6574 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 1 from Patent WO0234896.
ACCESSION AX429512
VERSION AX429512.1 GI:21540791
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kapeller-Libermann, R.
TITLE 13245, a novel human myotonic dystrophy type protein kinase and
uses therefor

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0234896-A 1 02-MAY-2002;

MILLENNIUM PHARM INC (US)

FEATURES

FEATURES	LOCATION/
SOURCE	1 6574

```
source
1. .65/4
/overwrite="Homocentri
```

```
/organism="Homo sapiens"
```

```
/mol_type="unassigned DN
```

ORIGIN

Query Match	100.0%	Score 6574	DB 6	Length 6574
<p> </p>				

Query Match	Score 65/4;	DB 6;	Length 65/4;
Best local similarity	100.0%	Score 65/4;	DB 6;
	100.0%	Score 65/4;	DB 6;
	100.0%	Score 65/4;	DB 6;

Best Local Similarity 100.0%; Pred. No. 0;

1	AGAGCGCGCAGTGGGGAGATGTTGAAGTTCAAATATATGAGCGCGGAATCCTTTGGATGCT	Qy
1	AGAGCGCGCAGTGGGGAGATGTTGAAGTTCAAATATATGAGCGCGGAATCCTTTGGATGCT	Db
60	AGAGCGCGCAGTGGGGAGATGTTGAAGTTCAAATATATGAGCGCGGAATCCTTTGGATGCT	Qy
60	AGAGCGCGCAGTGGGGAGATGTTGAAGTTCAAATATATGAGCGCGGAATCCTTTGGATGCT	Db
61	GGTGCTGCTGAACCCATTGCCAGCGCGGCCCTCCAGGCTGAATCTGTCTTCCTCCAGGGGAAA	Qy
61	GGTGCTGCTGAACCCATTGCCAGCGCGGCCCTCCAGGCTGAATCTGTCTTCCTCCAGGGGAAA	Db
121	CCACCCCTTATGACTCAACAGCAGATGTCCTCTCTTTCCGAGAGGGGATATTAGATGCC	Qy
121	CCACCCCTTATGACTCAACAGCAGATGTCCTCTCTTTCCGAGAGGGGATATTAGATGCC	Db
181	CTCTTTGTTCTCTTTGAAGAATGCAGTCAGCGCTCTCTGATGAAGATTAAGACACGTGAGC	Qy
181	CTCTTTGTTCTCTTTGAAGAATGCAGTCAGCGCTCTCTGATGAAGATTAAGACACGTGAGC	Db
241	AACTTTGTCCGGAAGTATTCGACACCATAGCTGAGTTACAGGAGTCCAGCCCTTCGGCA	Qy
241	AACTTTGTCCGGAAGTATTCGACACCATAGCTGAGTTACAGGAGTCCAGCCCTTCGGCA	Db
301	ANGGACTTTGGAAGTCAGAAAGCTTTGTAGGTGTGGTCACCTTCTCTGNAAGTGCAGGTTGTA	Qy
301	ANGGACTTTGGAAGTCAGAAAGCTTTGTAGGTGTGGTCACCTTCTCTGNAAGTGCAGGTTGTA	Db
361	AGAGAGAAGCAACCCGGGACATCTATGCTATGAAAGTGATGAAGAAGAAGGCTTTATTG	Qy
361	AGAGAGAAGCAACCCGGGACATCTATGCTATGAAAGTGATGAAGAAGAAGGCTTTATTG	Db
421	GCCCAAGAGCAGGTTTCAATTTTGTGAGAGAGCGGAACAATATATCTCGAAGCACAAAG	Qy
421	GCCCAAGAGCAGGTTTCAATTTTGTGAGAGAGCGGAACAATATATCTCGAAGCACAAAG	Db
481	CGTGCGATCCCCCAATTACAGTATGCGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG	Qy
481	CGTGCGATCCCCCAATTACAGTATGCGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG	Db
541	GAATATCAGCCTGGAGGGAATTGCTGTCATTTTGAATAGATATGAGGACCAAGTTAGAT	Qy
541	GAATATCAGCCTGGAGGGAATTGCTGTCATTTTGAATAGATATGAGGACCAAGTTAGAT	Db
601	GAACAACCTGTACAGTTTTTACTAGCTGAGCTGATTTTGCTGTTTCACAGCGTTCACTG	Qy
601	GAACAACCTGTACAGTTTTTACTAGCTGAGCTGATTTTGCTGTTTCACAGCGTTCACTG	Db
661	ATGGGATACGTGCATCGAGACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGGACAC	Qy
661	ATGGGATACGTGCATCGAGACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACAC	Db
721	ATCAGCTGTTGGATTTTGGATCTGCCGGAANAATGAATTCAAACAAGATGGTGAATGCC	Qy
721	ATCAGCTGTTGGATTTTGGATCTGCCGCGGAAAAATGAATTCAAACAAGATGGTGAATGCC	Db
781	AAACTCCCGATTGGGACCCAGATTACATGCTCCTCTGAACTGCTGATGAAGACCGG	Qy
781	AAACTCCCGATTGGGACCCAGATTACATGCTCCTCTGAACTGCTGATGAAGACCGG	Db
841	GATGAAAAGGCACTACGGCCTGGACTGTGATCTGTGGTTCAGTGGGCGTGAATGCCAT	Qy
841	GATGAAAAGGCACTACGGCCTGGACTGTGATCTGTGGTTCAGTGGGCGTGAATGCCAT	Db

1981 AGGGAGCTGGAGAACTCGAGAACCGAGAGATTCTTCTGAAGGCATCAGAAAGAGCTG 2040
2041 GTGGAAGCTGAGGAACGGCCGCAATCTCTCTGGAGAACAAAGGTAAAGAGACTAGAGACCATG 2100
2041 GTGGAAGCTGAGGAACGGCCGCAATCTCTCTGGAGAACAAAGGTAAAGAGACTAGAGACCATG 2100
2101 GAGCCTAGAGAAAACAGACTGAGAGATACATCCAGACAAAATCCCAACAGATCCAGCAG 2160
2101 GAGCCTAGAGAAAACAGACTGAGAGATACATCCAGACAAAATCCCAACAGATCCAGCAG 2160
2161 ATGGCTGATAAAATCTCTGAGCTCGAAGAGAAACATCGGAGGGCCCAAGTCTCAGCCCCAG 2220
2161 ATGGCTGATAAAATCTCTGAGCTCGAAGAGAAACATCGGAGGGCCCAAGTCTCAGCCCCAG 2220
2221 CACCTAGAGTGCACCTGAAACAGAAAGAGAGACGACTATGAGGAAAAGATTAAAGTGTG 2280
2221 CACCTAGAGTGCACCTGAAACAGAAAGAGAGAGACGACTATGAGGAAAAGATTAAAGTGTG 2280
2281 GACAAATCAGATAAAGAAAGACCTGGCTGACAAAGAGACACTGGAGAACATGATCAGAGA 2340
2281 GACAAATCAGATAAAGAAAGACCTGGCTGACAAAGAGACACTGGAGAACATGATGAGAGA 2340
2341 CACGAGGAGGCGCCATGAGAGGGCCAAAATTTCTCAGCGAACAGAAAGCGGATGATCAAT 2400
2341 CACGAGGAGGCGCCATGAGAGGGCCAAAATTTCTCAGCGAACAGAAAGCGGATGATCAAT 2400
2401 GCTATGGATTCCAGATCAGATCCCTGGAAACAGAGGATTGTGGACCTCTGAGGCCAAT 2460
2401 GCTATGGATTCCAGATCAGATCCCTGGAAACAGAGGATTGTGGACCTCTGAGGCCAAT 2460
2461 AAACCTTGACGAAATAGCAGTCTTTTATCCCAAGGAAACATGAAAGGCCCAAGAGAGATG 2520
2461 AAACCTTGACGAAATAGCAGTCTTTTATCCCAAGGAAACATGAAAGGCCCAAGAGAGATG 2520
2521 ATTTCTGAACCTCAGCGAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCC 2580
2521 ATTTCTGAACCTCAGCGAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCC 2580
2581 CAGAACCGAAACCTGGAGGAGCAGCTGGAGAGATCAGCCACCAAGACACAGTGACAAAG 2640
2581 CAGAACCGAAACCTGGAGGAGCAGCTGGAGAGATCAGCCACCAAGACACAGTGACAAAG 2640
2641 AATCGGCTGCTGGAACTGGAGAACAGATTGCGGAGGTCACTGTAGACACAGAGAGCAG 2700
2641 AATCGGCTGCTGGAACTGGAGAACAGATTGCGGAGGTCACTGTAGACACAGAGAGCAG 2700
2701 ARACTGGAGCTCAAGGCGCAGCTCAGAGCTACAGCTCTCCCTGCAGAGGCGGAGTCA 2760
2701 ARACTGGAGCTCAAGGCGCAGCTCAGAGCTACAGCTCTCCCTGCAGAGGCGGAGTCA 2760
2761 CAGTTTGACAGCCCTGCAGGCTGCAGGCGGCGCTGGAGAGCCAGCTTTCGCCAGGCGAAG 2820
2761 CAGTTTGACAGCCCTGCAGGCTGCAGGCGGCGCTGGAGAGCCAGCTTTCGCCAGGCGAAG 2820
2821 ACAGAGCTGGAGAGACACACAGAGCTGAAGAGAGATCCAGGCACTCAGGCGACAT 2880
2821 ACAGAGCTGGAGAGACACACAGAGCTGAAGAGAGATCCAGGCACTCAGGCGACAT 2880
2881 AGAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAAATCAGAGC 2940
2881 AGAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAAATCAGAGC 2940
2941 CTGGAGGAGCAGCTAAACAGCTGACCGAGGACAAACGCTGAACCTCAAACCAAACTTC 3000
2941 CTGGAGGAGCAGCTAAACAGCTGACCGAGGACAAACGCTGAACCTCAAACCAAACTTC 3000
3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGAGAGATTGTCAACTCGGA 3060
3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGAGAGATTGTCAACTCGGA 3060
3061 AGTGAAGTGGACCATCTCCGCGGAGATCAGGAAACAGAGATCAGCTTACAGCCAG 3120

3061 AGTGAAGTGGACCATCTCCGCGGAGATCACGAACGAGAGATGACGCTTACCAGCCAG 3120
3121 AAGCAAAACGATGAGGCTCTGAAGACACAGTGCACCATGCTGGAGGAAACAGTCTATGGAT 3180
3121 AAGCAAAACGATGAGGCTCTGAAGACACAGTGCACCATGCTGGAGGAAACAGTCTATGGAT 3180
3181 TTGAGGCGCTTAAACGATGAGCTGTAGAAAAGAGAGCGGAGGCGCTGGAGGAGC 3240
3181 TTGAGGCGCTTAAACGATGAGCTGTAGAAAAGAGAGCGGAGGCGCTGGAGGAGC 3240
3241 GTCTGGGTGATGAGAAATCCAGTTTGAGTGTGGGTTCGAGAGCTGACAGAAATGCTG 3300
3241 GTCTGGGTGATGAGAAATCCAGTTTGAGTGTGGGTTCGAGAGCTGACAGAAATGCTG 3300
3301 GACACGAGAAAACAGAGCAGGCGAGAGCCGATCAGCGGATCACCGAGTCTCCGAGGTG 3360
3301 GACACGAGAAAACAGAGCAGGCGAGAGCCGATCAGCGGATCACCGAGTCTCCGAGGTG 3360
3361 GTGAGCTGCGACTGAAGAGGACAAAGGCTGAGATTCTCGCTCTGCAGCAGGCTCTCAAA 3420
3361 GTGAGCTGCGACTGAAGAGGACAAAGGCTGAGATTCTCGCTCTGCAGCAGGCTCTCAAA 3420
3421 GAGCAGAGCTGAAGCGCGAGAGCCCTCTCTGCAAGCTCAATGACCTGGAGAAAGAGCAT 3480
3421 GAGCAGAGCTGAAGCGCGAGAGCCCTCTCTGCAAGCTCAATGACCTGGAGAAAGAGCAT 3480
3481 GCTATGCTTGAATGAATGCTGCCAAGCTTACACAGAGAGCTGAGAGCTGAAACGAGGCTC 3540
3481 GCTATGCTTGAATGAATGCTGCCAAGCTTACACAGAGAGCTGAGAGCTGAAACGAGGCTC 3540
3541 AAACAGAGGCTTCTGGAAGAGCAAGCAAAATTTACAGCAGCAGATGGAACCTGCGAGAAAAT 3600
3541 AAACAGAGGCTTCTGGAAGAGCAAGCAAAATTTACAGCAGCAGATGGAACCTGCGAGAAAAT 3600
3601 CACATTTTCGCTGACTCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTACTGAAG 3660
3601 CACATTTTCGCTGACTCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTACTGAAG 3660
3661 ACAGAAAGAGTCACTTGGAGTATCAGCTGGAAAAATTCAGGTCTCTATTCTCATGAA 3720
3661 ACAGAAAGAGTCACTTGGAGTATCAGCTGGAAAAATTCAGGTCTCTATTCTCATGAA 3720
3721 AAGGTGAAATGGAAGGCACTATTCTCAACAAACAAACTCATTTCTTTCGAGGCC 3780
3721 AAGGTGAAATGGAAGGCACTATTCTCAACAAACAAACTCATTTCTTTCGAGGCC 3780
3781 AAATGACCAACCTGCTAAAGAAAAAGGGTTTATTTAGTCGACGGAAGAGGACCCT 3840
3781 AAATGACCAACCTGCTAAAGAAAAAGGGTTTATTTAGTCGACGGAAGAGGACCCT 3840
3841 GCTTTACCCACACAGGTTCTCTGCAAGTCAATGAGCTGAAAGCTGGGAGAGGAG 3900
3841 GCTTTACCCACACAGGTTCTCTGCAAGTCAATGAGCTGAAAGCTGGGAGAGGAG 3900
3901 AAGCTCGCTGTCAGAGCTAGAGGAGCCCTTCAAGAACCCGATCGAGCTCCGGTCC 3960
3901 AAGCTCGCTGTCAGAGCTAGAGGAGCCCTTCAAGAACCCGATCGAGCTCCGGTCC 3960
3961 GCCCGGAGGAAGCTGCCACCGCAAAAGCAACGGAACCAACCCATCCAGCCAGGCC 4020
3961 GCCCGGAGGAAGCTGCCACCGCAAAAGCAACGGAACCAACCCATCCAGCCAGGCC 4020
4021 ACCGCGAGCAGAGATCGCCATGTCGCCATCGTGGGTGCGGAGCAGCAGCCAGT 4080
4021 ACCGCGAGCAGAGATCGCCATGTCGCCATCGTGGGTGCGGAGCAGCAGCCAGT 4080
4081 GCCATGAGCTGTGCGCCCGCCATCAGCGCGAGAAAAGAGTCTTCACTCCAGAGGAA 4140
4081 GCCATGAGCTGTGCGCCCGCCATCAGCGCGAGAAAAGAGTCTTCACTCCAGAGGAA 4140
4141 TTTAGTCGGGCTTTAAGGAAACGATGACCAATATTCTTCAACGATTTCAACGATAGGA 4200
4141 TTTAGTCGGGCTTTAAGGAAACGATGACCAATATTCTTCAACGATTTCAACGATAGGA 4200

4201 CTGAACATGGAGGCCACAAAGTGTGTGTCTGGATACCGTGCACCTTTGGAGCCGAG 4260
|||||
4201 CTGAACATGGAGGCCACAAAGTGTGTGTCTGGATACCGTGCACCTTTGGAGCCGAG 4260
|||||
4261 GCATCCAAATGTCTCGAATGTCAAGTGTGTCACCCCAAGTGTCCACGTCCTTGCCA 4320
|||||
4261 GCATCCAAATGTCTCGAATGTCAAGTGTGTCACCCCAAGTGTCCACGTCCTTGCCA 4320
|||||
4321 GCCACCTGCGGCTTGCTGTGTAATATGCAACACATTCACCGAGGCTTCTGCGGTGAC 4380
|||||
4321 GCCACCTGCGGCTTGCTGTGTAATATGCAACACATTCACCGAGGCTTCTGCGGTGAC 4380
|||||
4381 AANAATGAATCCCGAGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGCACTGGAAGG 4440
|||||
4381 AANAATGAATCCCGAGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGCACTGGAAGG 4440
|||||
4441 TGGATGAAGGTGCCCGAGGAATAACAAACGAGGACAGCAAGGCTGGACAGGAAGTACATT 4500
|||||
4441 TGGATGAAGGTGCCCGAGGAATAACAAACGAGGACAGCAAGGCTGGACAGGAAGTACATT 4500
|||||
4501 GTCTCGAGGGATCAAAAGTCTCATTTATGACATGAAAGCCAGAGAGCTGGACAGAG 4560
|||||
4501 GTCTCGAGGGATCAAAAGTCTCATTTATGACATGAAAGCCAGAGAGCTGGACAGAG 4560
|||||
4561 CCGGTGGAAGAAATTTGAGCTGTGCTTCCCGAGCGGGATGTATCTATTCAATGCTGCCGT 4620
|||||
4561 CCGGTGGAAGAAATTTGAGCTGTGCTTCCCGAGCGGGATGTATCTATTCAATGCTGCCGT 4620
|||||
4621 GGTGCTTCGAACTCCGAAATACAGCCAAAGCAGAAAGCAAGAGCTGATGCTTAACTG 4680
|||||
4621 GGTGCTTCGAACTCCGAAATACAGCCAAAGCAGAAAGCAAGAGCTGATGCTTAACTG 4680
|||||
4681 CTTGGAAATCCCTGCTGAACTGGAAAGTGTGACCGCTAGACATGAATCTGCACGCTG 4740
|||||
4681 CTTGGAAATCCCTGCTGAACTGGAAAGTGTGACCGCTAGACATGAATCTGCACGCTG 4740
|||||
4741 CCCTTCAGTGACAGGTGTGTTGGGACACGAGGAAGGCTCTACGCCCTGAAATGTC 4800
|||||
4741 CCCTTCAGTGACAGGTGTGTTGGGACACGAGGAAGGCTCTACGCCCTGAAATGTC 4800
|||||
4801 TTGAAATCCCTCTACCCATGTCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
|||||
4801 TTGAAATCCCTCTACCCATGTCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
|||||
4861 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAAAGCGGCACTGTGTCTTGTGGAC 4920
|||||
4861 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAAAGCGGCACTGTGTCTTGTGGAC 4920
|||||
4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4980
|||||
4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4980
|||||
4981 CCCAAATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5040
|||||
4981 CCCAAATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5040
|||||
5041 GGGCTCTGCATCTGTGGAGCCATGCCAGAAAGTGTGCTCATTTCTCGCTACAAAGAAAC 5100
|||||
5041 GGGCTCTGCATCTGTGGAGCCATGCCAGAAAGTGTGCTCATTTCTCGCTACAAAGAAAC 5100
|||||
5101 CTCAGCAATATCTGCAATCCGAAAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160
|||||
5101 CTCAGCAATATCTGCAATCCGAAAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160
|||||
5161 TTCACCAATTTACAGTATCTCATTTGAAACCAATAAATTTACGAAATCGACATGAAGCAG 5220
|||||
5161 TTCACCAATTTACAGTATCTCATTTGAAACCAATAAATTTACGAAATCGACATGAAGCAG 5220
|||||
5221 TACACGCTCGAGGAATTCCTGGATAGAAATGACATTCCTTGGCAGCTCTGCTGTTGCC 5280
|||||
5221 TACACGCTCGAGGAATTCCTGGATAGAAATGACATTCCTTGGCAGCTCTGCTGTTGCC 5280
|||||

5281 GCCTCTTCCACAGCTTCCCTGTCTCAATCTGTCAGGTGAACAGCGCAGGCGCAGAG 5340
|||||
5281 GCCTCTTCCACAGCTTCCCTGTCTCAATCTGTCAGGTGAACAGCGCAGGCGCAGAG 5340
|||||
5341 GAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTCTGTGATTTCTTACGGAACGCTAGC 5400
|||||
5341 GAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTCTGTGATTTCTTACGGAACGCTAGC 5400
|||||
5401 CGCAACAGCATCTCAAGTGGAGTCCCTTACCTTTGGCCTTTGCCCTACAGAGAACCTTAT 5460
|||||
5401 CGCAACAGCATCTCAAGTGGAGTCCCTTACCTTTGGCCTTTGCCCTACAGAGAACCTTAT 5460
|||||
5461 CTGTTTGTGACCCACTTCAACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 5520
|||||
5461 CTGTTTGTGACCCACTTCAACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 5520
|||||
5521 GGGACCCCTGCGGAGGCTACCTGGACATCCGAAACCCGCGCTACCTGGGCCCCGCAATT 5580
|||||
5521 GGGACCCCTGCGGAGGCTACCTGGACATCCGAAACCCGCGCTACCTGGGCCCCGCAATT 5580
|||||
5581 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAATTAAGGTCATTGCTGC 5640
|||||
5581 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAATTAAGGTCATTGCTGC 5640
|||||
5641 AAGGAAACCTCTGTGAAGGATCCGGCACTGAACAACACCCGGGGCCGCTCCACCTCCCGC 5700
|||||
5641 AAGGAAACCTCTGTGAAGGATCCGGCACTGAACAACACCCGGGGCCGCTCCACCTCCCGC 5700
|||||
5701 AGCAGCCCAACAGCAGGAGCCCAACCACTGACAAAGAGCAGATCAACAGCGCGTGCC 5760
|||||
5701 AGCAGCCCAACAGCAGGAGCCCAACCACTGACAAAGAGCAGATCAACAGCGCGTGCC 5760
|||||
5761 TCCAGCCAGCGCCCGCAGGAGCCCGAGCCACCCGAGAGCCAAAGCAACCCCAACCGC 5820
|||||
5761 TCCAGCCAGCGCCCGCAGGAGCCCGAGCCACCCGAGAGCCAAAGCAACCCCAACCGC 5820
|||||
5821 TACCGCAGGCGCGGAGCCGAGCTCCGAGGAGCAAGTCTCTCGCCGCCCTCGGAGCGA 5880
|||||
5821 TACCGCAGGCGCGGAGCCGAGCTCCGAGGAGCAAGTCTCTCGCCGCCCTCGGAGCGA 5880
|||||
5881 GAGAAGTCCCCCGCCCGATGCTCAGCACCGGAGAGCGGTCCCCCGGGAGGCTGTTT 5940
|||||
5881 GAGAAGTCCCCCGCCCGATGCTCAGCACCGGAGAGCGGTCCCCCGGGAGGCTGTTT 5940
|||||
5941 GAAGACAGCAGAGGCGCGCTGCTGCGGAGCGCTGAGGAGCCCGCTGTCCAGGTC 6000
|||||
5941 GAAGACAGCAGAGGCGCGCTGCTGCGGAGCGCTGAGGAGCCCGCTGTCCAGGTC 6000
|||||
6001 AACAGGAAAGAGGCGAGAGTCTCTCAAGTTTTCACGGTTTAACTGTCACTGTCACTATTAT 6060
|||||
6001 AACAGGAAAGAGGCGAGAGTCTCTCAAGTTTTCACGGTTTAACTGTCACTGTCACTATTAT 6060
|||||
6061 GACTGGAATAAAGCTGGACAACCTGCGCAGCTAACTGTGTCAGTCTGAGGATCATCCAG 6120
|||||
6061 GACTGGAATAAAGCTGGACAACCTGCGCAGCTAACTGTGTCAGTCTGAGGATCATCCAG 6120
|||||
6121 CTGAATCGAGAAATCCCGCAGCAGGTTGAAAGTCTGTTCTGAGAAACAGATTTTGTCTGA 6180
|||||
6121 CTGAATCGAGAAATCCCGCAGCAGGTTGAAAGTCTGTTCTGAGAAACAGATTTTGTCTGA 6180
|||||
6181 GCAGAGTTCATGTGACTTCTAGACGTGTCATTAATAATGCTTAAAGGCTTAAAGGCTGACAGC 6240
|||||
6181 GCAGAGTTCATGTGACTTCTAGACGTGTCATTAATAATGCTTAAAGGCTTAAAGGCTGACAGC 6240
|||||
6241 CAGCCACTCTGCTTACAAAAGAGTACTTATGAGTCACTGACGTGAAGAAACATTTGTA 6300
|||||
6241 CAGCCACTCTGCTTACAAAAGAGTACTTATGAGTCACTGACGTGAAGAAACATTTGTA 6300
|||||
6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTTCCCTGGAGCC 6360
|||||
6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTTCCCTGGAGCC 6360
|||||
6361 GAGAGACAACTGTGTGTTGATTTGAAGGACAGGACCAACACCTGATTTAGTTCCA 6420
|||||

Db 1501 TCCCTCTGGAGCAGGACCTTGTACTACTACATCACAGATCCAGTAGCTTAAGCGAAGT 1560
Qy 1579 TTGGAGCAAGCAGCGATGGAGGTGTCCAGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1638
Db 1561 TTGGAGCAAGCAGCGATGGAGGTGTCCAGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1620
Qy 1639 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCCAGGCTCAA 1698
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCCAGGCTCAA 1680
Qy 1699 GTGGAAGAAATGAGGTGATGATGAATCAATGATGGAAGAGGATCTTGTCTCAGCAAGAAGA 1758
Db 1681 GTGGAAGAAATGAGGTGATGATGAATCAATGATGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740
Qy 1759 CGGAGTATCTCTACGATCTGAGTCTGAGAGAGTCTCGGCTTCTGCTGGAAGAAATTCAG 1818
Db 1741 CGGAGTATCTCTACGATCTGAGTCTGAGAGAGTCTCGGCTTCTGCTGGAAGAAATTCAG 1800
Qy 1819 CGGAAAGCGACAGAAATGTTCAGCATAAATCTTGAAGGCTTAAGGATCAAGGAAAGCCTGAA 1878
Db 1801 CGGAAAGCGACAGAAATGTTCAGCATAAATCTTGAAGGCTTAAGGATCAAGGAAAGCCTGAA 1860
Qy 1879 GTGGAGAAATGCGAACTGGAGAAATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938
Db 1861 GTGGAGAAATGCGAACTGGAGAAATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
Qy 1939 CTCGAGAGAAATGCGAAAGGCTGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTG 1998
Db 1921 CTCGAGAGAAATGCGAAAGGCTGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTG 1980
Qy 1999 CAGAACCGAGAGGATCTTCTGAAGGATCAGAAAGAGCTGTGTGAAGCTGAGGAACGC 2058
Db 1981 CAGAACCGAGAGGATCTTCTGAAGGATCAGAAAGAGCTGTGTGAAGCTGAGGAACGC 2040
Qy 2059 CGCATTTCTTGGAGAACAGGTAAGAGACTAGAGACCATGAGCGTGAAGAAACAGA 2118
Db 2041 CGCATTTCTTGGAGAACAGGTAAGAGACTAGAGACCATGAGCGTGAAGAAACAGA 2100
Qy 2119 CTGAAGGATGACATCCAGACAAATCCACAGATCCAGCAGATGGCTGATAAATTTCTG 2178
Db 2101 CTGAAGGATGACATCCAGACAAATCCACAGATCCAGCAGATGGCTGATAAATTTCTG 2160
Qy 2179 GAGCTCGAGAGAAACATCGGAGGCCAGTCTCAGCCAGCAGCCTAGAGTGCACCTG 2238
Db 2161 GAGCTCGAGAGAAACATCGGAGGCCAGTCTCAGCCAGCAGCCTAGAGTGCACCTG 2220
Qy 2239 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAAGAAA 2298
Db 2221 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAAGAAA 2280
Qy 2299 GAGCTGGCTGACAGAGACACTGGAGAACATGATCGAGAGACAGCAGAGAGAGGCCCAT 2358
Db 2281 GAGCTGGCTGACAGAGACACTGGAGAACATGATCGAGAGACAGCAGAGAGAGGCCCAT 2340
Qy 2359 GAGAAAGGCAAAATTTCTCAGCGAACAGAGGCCATGATCAATGCTATGATTTCCAGATC 2418
Db 2341 GAGAAAGGCAAAATTTCTCAGCGAACAGAGGCCATGATCAATGCTATGATTTCCAGATC 2400
Qy 2419 AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTCAGCAAAATAGC 2478
Db 2401 AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTCAGCAAAATAGC 2460
Qy 2479 AGTCTTTTACCCAAAGGAACATGAGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAA 2538
Db 2461 AGTCTTTTACCCAAAGGAACATGAGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAA 2520
Qy 2539 CAGAAATTTTACCTGGAGACAGAGCTGGGAAAGTTGGAGGCCCAAGAACCGAAATTCGAG 2598
Db 2521 CAGAAATTTTACCTGGAGACAGAGCTGGGAAAGTTGGAGGCCCAAGAACCGAAATTCGAG 2580
Qy 2599 GAGCAGCTGGAGAGATCAGCCACCAAGACCAAGTGAAGAAATTCGCTGTGGAACCTG 2658
Db 2581 GAGCAGCTGGAGAGATCAGCCACCAAGACCAAGTGAAGAAATTCGCTGTGGAACCTG 2640

Qy 2659 GAGACAAAGATTCGGGAGAGTCAAGTCTAGAGCAAGAGGAGCAAACTGGAGCTCAAGGCG 2718
Db 2641 GAGACAAAGATTCGGGAGAGTCAAGTCTAGAGCAAGAGGAGCAAACTGGAGCTCAAGGCG 2700
Qy 2719 CAGCTTCAGAGCTACAGCTCTCCCTGAGAGCGCGAGTCAAGTTCAGAGCCCTGAG 2778
Db 2701 CAGCTTCAGAGCTACAGCTCTCCCTGAGAGCGCGAGTCAAGTTCAGAGCCCTGAG 2760
Qy 2779 GCTGACCGGGGGCCCTGGAGAGCAGCTTCGCCAGCGCAAGAGACAGAGCTGGAAGAGACC 2838
Db 2761 GCTGACCGGGGGCCCTGGAGAGCAGCTTCGCCAGCGCAAGAGACAGAGCTGGAAGAGACC 2820
Qy 2839 ACAGCAGAGCTGAAGAGAGATCCAGGCATCCAGGCACATAGAGATGAATTCAGGCG 2898
Db 2821 ACAGCAGAGCTGAAGAGAGATCCAGGCATCCAGGCACATAGAGATGAATTCAGGCG 2880
Qy 2899 AAATTTGATGCTCTTCGTAAACAGAGTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2958
Db 2881 AAATTTGATGCTCTTCGTAAACAGAGTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2940
Qy 2959 CAGCTGACCGAGGACACGCTGAACTCAACCAACAACTTCTACTTGTCCAAACAACTC 3018
Db 2941 CAGCTGACCGAGGACACGCTGAACTCAACCAACAACTTCTACTTGTCCAAACAACTC 3000
Qy 3019 GATGAGGCTTCTGGCGCCACAGCAGATTTGTAACCTGCGAAGTGAAGTGGACCATCTC 3078
Db 3001 GATGAGGCTTCTGGCGCCACAGCAGATTTGTAACCTGCGAAGTGAAGTGGACCATCTC 3060
Qy 3079 CGCCGGAGATCAACGGAACGAGAGATCCAGCTTACCAGCCAGAAACAAACGATGGAGCT 3138
Db 3061 CGCCGGAGATCAACGGAACGAGAGATCCAGCTTACCAGCCAGAAACAAACGATGGAGCT 3120
Qy 3139 CTGAAGACCACTGCACCATCTCGAGGAAACAGCTCATGATTTGGAGGCCCTTAAACGAT 3198
Db 3121 CTGAAGACCACTGCACCATCTCGAGGAAACAGCTCATGATTTGGAGGCCCTTAAACGAT 3180
Qy 3199 GAGCTGTAGAAAAAGAGCGCAGTGGAGGCCCTGGAGGAGCTCTGGGTGATGAGAAA 3258
Db 3181 GAGCTGTAGAAAAAGAGCGCAGTGGAGGCCCTGGAGGAGCTCTGGGTGATGAGAAA 3240
Qy 3259 TCCAGTTTGAGTGTCCGGTTTCAGAGCTGAGAGATGCTGCACACCCAGAAACAGAGC 3318
Db 3241 TCCAGTTTGAGTGTCCGGTTTCAGAGCTGAGAGATGCTGCACACCCAGAAACAGAGC 3300
Qy 3319 AGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCCAGCTGGTGGAGCTGGCAGTGAAG 3378
Db 3301 AGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCCAGCTGGTGGAGCTGGCAGTGAAG 3360
Qy 3379 GAGCAACAGGCTGAGATTTCTGCTCTCAGAGGCTCTCAAGAGCAGAGCTGAAGGCC 3438
Db 3361 GAGCAACAGGCTGAGATTTCTGCTCTCAGAGGCTCTCAAGAGCAGAGCTGAAGGCC 3420
Qy 3439 GAGAGCTCTCTGACAAAGCTCAATGACCTGAGAGAGCATGCTATGCTTGAATGAAT 3498
Db 3421 GAGAGCTCTCTGACAAAGCTCAATGACCTGAGAGAGCATGCTATGCTTGAATGAAT 3480
Qy 3499 GCCGAAAGCTTACAGCAGAGCTGGAGACTGAAACGAGCTCAAAACAGAGGCTTCTGGAA 3558
Db 3481 GCCGAAAGCTTACAGCAGAGCTGGAGACTGAAACGAGCTCAAAACAGAGGCTTCTGGAA 3540
Qy 3559 GAGCAACGCAAAATACAGCAGAGCTGGAGCTCCAGAAATCACATTTTCGCTGACT 3618
Db 3541 GAGCAACGCAAAATACAGCAGAGCTGGAGCTCCAGAAATCACATTTTCGCTGACT 3600
Qy 3619 CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3678
Db 3601 CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660
Qy 3679 GAGTATCAGCTGGAAGAACTTCAAGTCTCTATCTCATGAAGGCTGAAGTGGAGGCTG 3738
Db 3661 GAGTATCAGCTGGAAGAACTTCAAGTCTCTATCTCATGAAGGCTGAAGTGGAGGCTG 3720

Db 5881 ATGCTCAGCAGCGGAGAGAGCGGTCCCGGGAGGCTGTTTGAAGACAGCAGCAGGGCG 5940
QY 5959 CGGTGCTCGGGAGCGGTGAGGACCCCGTGTCCAGGTGAACAAGGGAAGAGGCGAG 6018
Db 5941 CGGTGCTCGGGAGCGGTGAGGACCCCGTGTCCAGGTGAACAAGGGAAGAGGCGAG 6000
QY 6019 AGTCCCTCAAGTTTTCACGGTTTAAACACTGTCACTTATTATGACTGGATATAAAGCTG 6078
Db 6001 AGTCCCTCAAGTTTTCACGGTTTAAACACTGTCACTTATTATGACTGGATATAAAGCTG 6060
QY 6079 GACAACCTGCCAGCTAACTGCTCAGTCTGTAGGATCATCCAGCTGAATGAGAAATCCGG 6138
Db 6061 GACAACCTGCCAGCTAACTGCTCAGTCTGTAGGATCATCCAGCTGAATGAGAAATCCGG 6120
QY 6139 CAGCAGGTGAAAGTCTGTTCTGAGAACAGATTATTGC 6177
Db 6121 CAGCAGGTGAAAGTCTGTTCTGAGAACAGATTATTGC 6159

RESULT 3
LOCUS AX504254 6298 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 43 from Patent WO0233099.
ACCESSION AX504254
VERSION AX504254.1 GI:23386094
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Gururajan, R., Baughn, M.R., Walla, N.K., Elliott, V.S., Xu, Y., Arvizu, C., Yao M.G., Ramkumar, J., Ding, L., Tang, Y.T., Hafalia, A.J., Nguyen, D.B., Gandhi, A.R., Lu, Y., Yue, H., Burford, N., Bandman, O., Triboley, C.M., Lal, P.G., Recipon, S.A., Lu, D.A., Borowsky, M.L., Thornton, M., Swannaker, A., Thangavelu, K., Khan, F.A. and Ison, C.H.
TITLE Human kinases
JOURNAL Patent: WO 0233099-A 43 25-APR-2002;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..6298
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7484498CB1"

ORIGIN
Query March 86.5%; Score 5684.4; DB 6; Length 6298;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5979; Conservative 0; Mismatches 36; Indels 243; Gaps 3;
QY 1 AGAGCGCCAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCT 60
Db 37 AGAGCGCCAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCT 96
QY 61 GGTGCTGCTGAACCCATTCGACCGCGGCTCAGCTGAATCTGTTTCCAGGGGAAA 120
Db 97 GGTGCTGCTGAACCCATTCGACCGCGGCTCAGCTGAATCTGTTTCCAGGGGAAA 156
QY 121 CCACCTTTATGACTCAACAGAGATGCTCCTCTTTCCGAGAGAGGATATTAGATGCC 180
Db 157 CCACCTTTATGACTCAACAGAGATGCTCCTCTTTCCGAGAGAGGATATTAGATGCC 216
QY 181 CTCTTTGTTCTCTTTGAAGATGACGTACGCTGCTCTGATGAAGATTAAGCAGCTGAGC 240
Db 217 CTCTTTGTTCTCTTTGAAGATGACGTACGCTGCTCTGATGAAGATTAAGCAGCTGAGC 276
QY 241 AACTTTGTCCGGAAGTATTCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCA 300
Db 277 AACTTTGTCCGGAAGTATTCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCA 336
QY 301 AAGGACTTCAAGTCAAGAGTCTTGTAGGTTGCTGCTCACTTTGCTGAAGTGCAGTGTTA 360

Db 337 AAGGACTTTCGAAGTCAGAAGTCTTGTAGGTTGTGCTCACTTTGTGTAAGTGCAGGTGTTA 396
QY 361 AGAGAGAAGCAACCGGGGACATCTATGCTATGAAGTGATGAAGAAGAGGCTTTATTG 420
Db 397 AGAGAGAAGCAACCGGGGACATCTATGCTATGAAGTGATGAAGAAGAGGCTTTATTG 456
QY 421 GCCAGGAGCAGGTTTCAATTTTGGAGAAAGAGCGGAAACATATTATCTCGAAGCACAAGC 480
Db 457 GCCAGGAGCAGGTTTCAATTTTGGAGAAAGAGCGGAAACATATTATCTCGAAGCACAAGC 516
QY 481 CGGTGGATCCCGCAATTCAGTATGCTTTTCAGACACAAATACCTTTATCTGGTCATG 540
Db 517 CGGTGGATCCCGCAATTCAGTATGCTTTTCAGACACAAATACCTTTATCTGGTCATG 576
QY 541 GAATATCAGCCTGGAGGGGACTTGTCTGCTCACTTTTGAATAGATATGAGGACCACTTAGAT 600
Db 577 GAATATCAGCCTGGAGGGGACTTGTCTGCTCACTTTTGAATAGATATGAGGACCACTTAGAT 636
QY 601 GAAACCTGTATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCAGAGGTTCACTCTG 660
Db 637 GAAACCTGTATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCAGAGGTTCACTCTG 696
QY 661 ATGGGATACGTGTCATCCGAGACATCAAGCCTGAGAACATCTCTGTTGACCGCACAGGACAC 720
Db 697 ATGGGATACGTGTCATCCGAGACATCAAGCCTGAGAACATCTCTGTTGACCGCACAGGACAC 756
QY 721 ATCAAGCTGTGGATTTTGGATCTGCCGCGGAAATGAATTCAAACAAGATGGTGAATGCC 780
Db 757 ATCAAGCTGTGGATTTTGGATCTGCCGCGGAAATGAATTCAAACAAGATGGTGAATGCC 816
QY 781 AAATCCCGATTGGGACCCCGAGATTACATGCTCCTCAAGTGCTGACTGTGATGAACCGG 840
Db 817 AAATCCCGATTGGGACCCCGAGATTACATGCTCCTCAAGTGCTGACTGTGATGAACCGG 876
QY 841 GATCGAAAGACCTACCGCTGAGCTGACTGTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
Db 877 GATCGAAAGACCTACCGCTGAGCTGACTGTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 936
QY 901 GAGATGATTTATGGAGATCCCTTCGACAGAGGACCTCTGCCAGAACCTTCAATATAC 960
Db 937 GAGATGATTTATGGAGATCCCTTCGACAGAGGAACTCTGCCAGAACCTTCAATATAC 996
QY 961 ATTATGAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTGCAGTGAATTT 1020
Db 997 ATTATGAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTGCAGTGAATTT 1056
QY 1021 CTTGATCTGATTCAGACTTGTGTGGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTT 1080
Db 1057 CTTGATCTGATTCAGACTTGTGTGGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTT 1116
QY 1081 TGCTGCCATCTTTCTCTCTAAATTTGACTGGAACAACTTCTGTAATCTCTCTCCCTCC 1140
Db 1117 TGCTGCCATCTTTCTCTCTAAATTTGACTGGAACAACTTCTGTAATCTCTCTCCCTCC 1176
QY 1141 TTGCTTCCCAACCTTCAAGTCTGACATGACACCTTCAATTTTGTATGAACACAGAGAAT 1200
Db 1177 TTGCTTCCCAACCTTCAAGTCTGACATGACACCTTCAATTTTGTATGAACACAGAGAAT 1236
QY 1201 TGCTGGGTTTCACTCTCTCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTG 1260
Db 1237 TGCTGGGTTTCACTCTCTCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTG 1296
QY 1261 CCGTTTGTGGGTTTTCGTACAGCAAGGCACTGGGATTTCTGGTAGATCTGTAGTCTGTT 1320
Db 1297 CCGTTTGTGGGTTTTCGTACAGCAAGGCACTGGGATTTCTGGTAGATCTGTAGTCTGTT 1356
QY 1321 GTGTGGGTTCTGACTCCCTCCAGACTAGTCTCCATGGAAAGAACTTCTCATCAA 1380
Db 1357 GTGTGGGTTCTGACTCCCTCCAGACTAGTCTCCATGGAAAGAACTTCTCATCAA 1416
QY 1381 AGCAAGAGCTACAAGACTCTCAGGACAAAGTGTCACAAGATGAGCAGGAAATGACCCG 1440
Db 1417 AGCAAGAGCTACAAGACTCTCAGGACAAAGTGTCACAAGATGAGCAGGAAATGACCCG 1476

QY 1441 TTACATCGGAGAGTGTACAGAGTGGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAG 1500
DB 1477 TTACATCGGAGAGTGTACAGAGTGGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAG 1536
QY 1501 GCCTCTGAGACTCAGAGATCCCTCTCTGGAGCAGGACCTTGTACCTAATCAGAGAAATGC 1560
DB 1537 GCCTCTGAGACTCAGAGATCCCTCTCTGGAGCAGGACCTTGTACCTAATCAGAGAAATGC 1596
QY 1561 AGTAGCTTAAGCGGAAGTTTGAGCAAGCAGCGATGGAGGTCTCCAGGAGGATCAGAAA 1620
DB 1597 AGTAGCTTAAGCGGAAGTTTGAGCAAGCAGCGATGGAGGTCTCCAGGAGGATCAGAAA 1656
QY 1621 GCAGTGCAGCTTCTCCATGATATCAGAGACAGAGCCGGAAGCTCCAGAGAAATCAAGAG 1680
DB 1657 GCAGTGCAGCTTCTCCATGATATCAGAGACAGAGCCGGAAGCTCCAGAGAAATCAAGAG 1716
QY 1681 CAGGAGTACAGGCTCAAGTGAAGAAATGAGGTGTGATGATCAATCAGTTGGAAGAGAT 1740
DB 1717 CAGGAGTACAGGCTCAAGTGAAGAAATGAGGTGTGATGATCAATCAGTTGGAAGAGAT 1776
QY 1741 CTTGTCTCAGCAAGAGACGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTT 1800
DB 1777 CTTGTCTCAGCAAGAGACGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTT 1836
QY 1801 GTGTCTGAAGAAATCAAGCGGAAACGGAAGAGATGTACGATTAATCTGTTGAAGCTTAAG 1860
DB 1837 GTGTCTGAAGAAATCAAGCGGAAACGGAAGAGATGTACGATTAATCTGTTGAAGCTTAAG 1896
QY 1861 GATCAGGGAAGCTGAGTGGAGAAATATGCGAAACTGAGAAATCAATCTCAGAGAG 1920
DB 1897 GATCAGGGAAGCTGAGTGGAGAAATATGCGAAACTGAGAAATCAATCTCAGAGAG 1956
QY 1921 CAGCTCAAAATTCAGGAGCTCCAGAGAAATCTGGAAGAGCTT----- 1962
DB 1957 CAGCTCAAAATTCAGGAGCTCCAGAGAAATCTGGAAGAGCTTGAAGAGCCGAGCAGGAG 2016
QY 1963 -----GCAAGGACGAGCGAGGAGGAGCTGGAG 1992
DB 2017 GGCACCGAGCTGTGCAGAAATTCGCCAGGCAAGAGGAGCGAGGAGGAGCTGGAG 2076
QY 1993 AAGCTGCAGAACCGAGAGGATCTCTGTAAGGATCAGAAAGAGCTGTGTAAGAGCTGAG 2052
DB 2077 AAGCTGCAGAACCGAGAGGATCTCTGTAAGGATCAGAAAGAGCTGTGTAAGAGCTGAG 2136
QY 2053 GAACGCGGCAATCTCTGAGAAACAAAGGTAAAGAGACTAGAGACCATGAGGAGCTAGAGAA 2112
DB 2137 GAACGCGGCAATCTCTGAGAAACAAAGGTAAAGAGACTAGAGACCATGAGGAGCTAGAGAA 2196
QY 2113 AACAGACTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATAAA 2172
DB 2197 AACAGACTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATAAA 2256
QY 2173 ATTCTGGAGCTCGAAGAGAAACATCCGGAGGCCAAGTCTCAGCCAGACCATAGAAAGTG 2232
DB 2257 ATTCTGGAGCTCGAAGAGAAACATCCGGAGGCCAAGTCTCAGCCAGACCATAGAAAGTG 2316
QY 2233 CAGCTGAAGAGAGGAGCAGCTATGAGGAAGATTAAGGTGTGACACATCAGATA 2292
DB 2317 CAGCTGAAGAGAGGAGCAGCTATGAGGAAGATTAAGGTGTGACACATCAGATA 2376
QY 2293 AAGAAAGACCTGTGCTGACAGGAGACACTGGAGAACATGATGTCAGAGACACAGAGAGGAG 2352
DB 2377 AAGAAAGACCTGTGCTGACAGGAGACACTGGAGAACATGATGTCAGAGACACAGAGAGGAG 2436
QY 2353 GCCCATGAGAGGCGAAAATCTCAGGAAACAGAGGCGATGATCAATGCTATGATTC 2412
DB 2437 GCCCATGAGAGGCGAAAATCTCAGGAAACAGAGGCGATGATCAATGCTATGATTC 2496
QY 2413 AAGATCAGATCCCTGGAACAGAGGATTTGTGGAATCTGTTGAAGCCAAATAACTTGCAGCA 2472
DB 2497 AAGATCAGATCCCTGGAACAGAGGATTTGTGGAATCTGTTGAAGCCAAATAACTTGCAGCA 2556

QY 2473 AATAGCAGTCTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATC 2532
DB 2557 AATAGCAGTCTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATC 2616
QY 2533 AGGCAACAGAAAATTTTACCTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGAACCGAAA 2592
DB 2617 AGGCAACAGAAAATTTTACCTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGAACCGAAA 2676
QY 2593 CTGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACAGTGCACAGAAATCGGCTGCTG 2652
DB 2677 CTGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACAGTGCACAGAAATCGGCTGCTG 2736
QY 2653 GAATCTGAGACAAAGATTGCGGAGGTGAGTCTAGAGCACGAGGACGAGAACTGGAGCTC 2712
DB 2737 GAATCTGAGACAAAGATTGCGGAGGTGAGTCTAGAGCACGAGGACGAGAACTGGAGCTC 2796
QY 2713 AAGCGCCAGCTCAGAGAGCTACAGTCTCTCCCTGCAAGAGCGCGAGTCAACAGTTGACAGCC 2772
DB 2797 AAGCGCCAGCTCAGAGAGCTACAGTCTCTCCCTGCAAGAGCGCGAGTCAACAGTTGACAGCC 2856
QY 2773 CTGAGGCTGCAAGCGCGGCTGGAGAGCAGCTTCGCCAGCGGAAAGACAGAGCTGGAA 2832
DB 2857 CTGAGGCTGCAAGCGCGGCTGGAGAGCAGCTTCGCCAGCGGAAAGACAGAGCTGGAA 2916
QY 2833 GAGACCAACAGCAAGCTGAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAATC 2892
DB 2917 GAGACCAACAGCAAGCTGAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAATC 2976
QY 2893 CAGCGCAAAATTTGATGCTCTCTGTTAAACAGCTGTAATCAACAGACCTGGAGAGCAG 2952
DB 2977 CAGCGCAAAATTTGATGCTCTCTGTTAAACAGCTGTAATCAACAGACCTGGAGAGCAG 3036
QY 2953 CTAACACAGCTGACCGAGGACAGCTGAATCAACCAACCAAACTTCTACTTGTCCAAA 3012
DB 3037 CTAACACAGCTGACCGAGGACAGCTGAATCAACCAACCAAACTTCTACTTGTCCAAA 3096
QY 3013 CAATCTGATGAGGCTTCTGCGCCCAACGACGAGATTTACAACCTGCGAAAGTGAAGTGAC 3072
DB 3097 CAATCTGATGAGGCTTCTGCGCCCAACGACGAGATTTACAACCTGCGAAAGTGAAGTGAC 3156
QY 3073 CATCTCCGCGGAGATCAGGAAACGAGAGATGAGCTTACCAGCCAGAGACCAACGATG 3132
DB 3157 CATCTCCGCGGAGATCAGGAAACGAGAGATGAGCTTACCAGCCAGAGACCAACGATG 3216
QY 3133 GAGCTCTGAAGACCACTGCGCCATCTGGAGGAAACAGGTCTATGGATTTGGAGGCCCTA 3192
DB 3217 GAGCTCTGAAGACCACTGCGCCATCTGGAGGAAACAGGTCTATGGATTTGGAGGCCCTA 3276
QY 3193 AACGATGAGCTGTAGAAAAGAGCGGACGTCAGGAGGCTGGAGGAGCGCTCTGGGTGAT 3252
DB 3277 AACGATGAGCTGTAGAAAAGAGCGGACGTCAGGAGGCTGGAGGAGCGCTCTGGGTGAT 3336
QY 3253 GAGAAATCCAGTTTGAGTGTGCGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAA 3312
DB 3337 GAGAAATCCAGTTTGAGTGTGCGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAA 3396
QY 3313 CAGAGCAGGCGGAGAGCGGATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCA 3372
DB 3397 CAGAGCAGGCGGAGAGCGGATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCA 3456
QY 3373 GTGAAGAGACAAAGGCTGAGATTTCTGCTCTGAGAGGCTCTCAAGAGCAGAACTG 3432
DB 3457 GTGAAGAGACAAAGGCTGAGATTTCTGCTCTGAGAGGCTCTCAAGAGCAGAACTG 3516
QY 3433 AAGCCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTTGA 3492
DB 3517 AAGCCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTTGA 3576
QY 3493 ATGAATGCCGAAGCTTACAGCAGAAAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTT 3552
DB 3577 ATGAATGCCGAAGCTTACAGCAGAAAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTT 3636
QY 3553 CTGGAAGAGCAGCCAAATTTACACGACAGATGGACCTGCAGAGAAAATCACATTTCCGT 3612

1261 TACAGCAAGGCACTGGGGAATCTTGGTAGATCTGAGTCTGTGTGTCGGGCTCTGGACTCC 1320
1339 CTTGCCAAGACTAGCTCCATCGGAAAGAACTTCTCATCAAAAAGCAAGAGCTACAGAC 1398
1321 CTTGCCAAGACTAGCTCCATCGGAAAGAACTTCTCATCAAAAAGCAAGAGCTACAGAC 1380
1399 TCTCAGCACAAGTGTACAGAGTGGACAGGAATGACCCGGTTACATCGGAGTGTCA 1458
1381 TCTCAGCACAAGTGTACAGAGTGGACAGGAATGACCCGGTTACATCGGAGTGTCA 1440
1459 GAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1518
1441 GAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1500
1519 TCCCTCTTGAGCAGGACCTTGTACTATCATCAGAAATGCAATAGCTTAAGCGAAGT 1578
1501 TCCCTCTTGAGCAGGACCTTGTACTATCATCAGAAATGCAATAGCTTAAGCGAAGT 1560
1579 TTGAGCAAGCAGGATGGAGGTGTCAGAGAGTGCACAAAGCACTGCAGCTTCTCCAT 1638
1561 TTGAGCAAGCAGGATGGAGGTGTCAGAGAGTGCACAAAGCACTGCAGCTTCTCCAT 1620
1639 GATATCAGAGCAGGACCGGAGTCTCCAGAAATCAAAGAGCAGGAGTACAGGCTCAA 1698
1621 GATATCAGAGCAGGACCGGAGTCTCCAGAAATCAAAGAGCAGGAGTACAGGCTCAA 1680
1699 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1759 CGGAGTCACTCTCAAGAACTGAGCTGAGAGAGTCTCGGCTGCTGCTCAAGAAATTCAG 1818
1741 CGGAGTCACTCTCAAGAACTGAGCTGAGAGAGTCTCGGCTGCTGCTCAAGAAATTCAG 1800
1819 CGGAAAGCAGAGTCTCAGCACTAACTGTTGAAAGCTTAAGATCAAGGGAGCTGGA 1878
1801 CGGAAAGCAGAGTCTCAGCACTAACTGTTGAAAGCTTAAGATCAAGGGAGCTGGA 1860
1879 GTGGGAAATATCGAACTGGAGAGAGTCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938
1861 GTGGGAAATATCGGAACTGGAGAGAGTCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
1939 CTCAGAGAAACTGGAGAGGCTG----- 1963
1921 CTCAGAGAAACTGGAGAGGCTGTAAGCCAGCAGGAGCCACCGAGCTGTGTCAG 1980
1964 -----CAAAGAGCAGCGGAGGAGCTGGAGAGCTGCAGAACTGCAGAACCGAGAG 2010
1981 AATATCCGCCAGGCAAGAGGAGCGAGCGGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2040
2011 GATCTTCTGAAAGGCACTCAGAAAGAGCTGTTGGAGAGCTGAGAAACCGCCCATCTCTG 2070
2041 GATCTTCTGAAAGGCACTCAGAAAGAGCTGTTGGAGAGCTGAGAAACCGCCCATCTCTG 2100
2071 GAGAAACAAGTAAAGAGCTAGAGACCATGGAGCTAGAGAAACAGACTGAAGATGAC 2130
2101 GAGAAACAAGTAAAGAGCTAGAGACCATGGAGCTAGAGAAACAGACTGAAGATGAC 2160
2131 ATCCAGACAAAATCCCAACAGATCCAGCAGAGTGGCTGATATAATTTCTGGAGCTCAGAG 2190
2161 ATCCAGACAAAATCCCAACAGATCCAGCAGAGTGGCTGATATAATTTCTGGAGCTCAGAG 2220
2191 AATCATCGGAGGCCCAAGCTCTCAGCCGACCTAGAGAGTGCCTGTAACAGAAAGAG 2250
2221 AATCATCGGAGGCCCAAGCTCTCAGCCGACCTAGAGAGTGCCTGTAACAGAAAGAG 2280
2251 CAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATTAAGAAAGACCTGGCTGAC 2310
2281 CAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATTAAGAAAGACCTGGCTGAC 2340
2311 AAGGACACTGAGAACATGATGAGACACAGGAGGAGGCCCATGAGAGGCCAA 2370

2341 AAGGACACTGAGAACATGATCAGAGACACAGGAGGAGGCCCATGAGAGGCCAA 2400
2371 ATTCTCAGCGAACAGAGGCGGATGATCAATCTATGATTTCCAGAGTCAGATCCCTCGAA 2430
2401 ATTCTCAGCGAACAGAGGCGGATGATCAATCTATGATTTCCAGAGTCAGATCCCTCGAA 2460
2431 CAGAGGATTGTGGAACTGTCTGAAAGCAATAAATTCAGCAAAATAGCAGTCTTTTACC 2490
2461 CAGAGGATTGTGGAACTGTCTGAAAGCAATAAATTCAGCAAAATAGCAGTCTTTTACC 2520
2491 CAAAGCAATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550
2521 CAAAGCAATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
2551 CTGAGACACAGGCTGGAGGCTTGGAGGCCAGAACCGAAACTGAGAGGAGCAGCTGGAG 2610
2581 CTGAGACACAGGCTGGAGGCTTGGAGGCCAGAACCGAAACTGAGAGGAGCAGCTGGAG 2640
2611 AAGATCAAGCAACCAAGACCACAGTGACAAGAACTCGGCTGCTGGAATCTGGAGACAAGATTG 2670
2641 AAGATCAAGCAACCAAGACCACAGTGACAAGAACTCGGCTGCTGGAATCTGGAGACAAGATTG 2700
2671 CGGAGGTCAGTCTAGAGCAGCAGGAGCAGAACTGAGGCTCAAGCGCCAGCTCACAGAG 2730
2701 CGGAGGTCAGTCTAGAGCAGCAGGAGCAGAACTGAGGCTCAAGCGCCAGCTCACAGAG 2760
2731 CTACAGCTCTCCCTGCGAGGAGCGGAGTCAAGTTGACAGCCCTGCGAGGCTGACCGGCG 2790
2761 CTACAGCTCTCCCTGCGAGGAGCGGAGTCAAGTTGACAGCCCTGCGAGGCTGACCGGCG 2820
2791 GCCCTGAGAGCAGCTTCCGAGGCGAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2850
2821 GCCCTGAGAGCAGCTTCCGAGGCGAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2880
2851 GAAGAGGAGTCCAGGCACTCAAGGACATAGAGTGAATTCAGCGCAAAATTTGATGCT 2910
2881 GAAGAGGAGTCCAGGCACTCAAGGACATAGAGTGAATTCAGCGCAAAATTTGATGCT 2940
2911 CTTCTGAAAGCTGATCTGTAATCAGAGCTGGAGAGCAGCTAAACAGCTGACCGAG 2970
2941 CTTCTGAAAGCTGATCTGTAATCAGAGCTGGAGAGCAGCTAAACAGCTGACCGAG 3000
2971 GACAACTGAACTCAACACCAAACTTCTACTGTCGCAAACTCCGATGAGGCTTCT 3030
3001 GACAACTGAACTCAACACCAAACTTCTACTGTCGCAAACTCCGATGAGGCTTCT 3060
3031 GGCCTCCAAACGAGATTTGACAACTCGAAGTGAAGTGGACCATCTCCGCGGAGATC 3090
3061 GGCCTCCAAACGAGATTTGACAACTCGAAGTGAAGTGGACCATCTCCGCGGAGATC 3120
3091 ACGGAAACAGAGATGACAGCTTACAGCAGAGCAAGCAAGTGGAGGCTCTGAGAGCCAG 3150
3121 ACGGAAACAGAGATGACAGCTTACAGCAGAGCAAGCAAGTGGAGGCTCTGAGAGCCAG 3180
3151 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTAAACAGTGAAGTCTAGAA 3210
3181 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTAAACAGTGAAGTCTAGAA 3240
3211 AAAGAGCGGAGTGGAGGCTTGGAGGAGCTTCTGGTGTATGAGAAATCCGAGTTTGA 3270
3241 AAAGAGCGGAGTGGAGGCTTGGAGGAGCTTCTGGTGTATGAGAAATCCGAGTTTGA 3300
3271 TGTGGGTTGAGAGCTGAGAGATGCTGGAACACCGAGAAACAGAGCAGGCGGAGAGCC 3330
3301 TGTGGGTTGAGAGCTGAGAGATGCTGGAACACCGAGAAACAGAGCAGGCGGAGAGCC 3360
3331 GATCAGCGGATCACCGAGTCTCGCAGGCTGGAGCTTCCGATGAGAGAGCAAGGCT 3390
3361 GATCAGCGGATCACCGAGTCTCGCAGGCTGGAGCTTCCGATGAGAGAGCAAGGCT 3420
3391 GAGATTCTCCCTCTGAGCAGGCTTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3450
3421 GAGATTCTCCCTCTGAGCAGGCTTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3480

QY 3451 GACAGCTCAATGAACTGGAGAAAGCATGTATGCTTTGAAATGAATGCCCGAAGCTTA 3510
Db 3481 GACAGCTCAATGAACTGGAGAAAGCATGTATGCTTTGAAATGAATGCCCGAAGCTTA 3540
QY 3511 CAGCAGAGCTGGAGACTGAAAGAGCTCAACAGAGCTTCTGGAAGAGCAAGCCAAA 3570
Db 3541 CAGCAGAGCTGGAGACTGAAAGAGCTCAACAGAGCTTCTGGAAGAGCAAGCCAAA 3600
QY 3571 TTACAGCAGCAGATGGACCTGACAGAAAAATCAATTTCCGTCTGACTCAAGGACTGCAA 3630
Db 3601 TTACAGCAGCAGATGGACCTGACAGAAAAATCAATTTCCGTCTGACTCAAGGACTGCAA 3660
QY 3631 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATGAGTATCAGCTG 3690
Db 3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATGAGTATCAGCTG 3720
QY 3691 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAAAAAGGAGCACTATTTCTCAA 3750
Db 3721 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAAAAAGGAGCACTATTTCTCAA 3780
QY 3751 CAAACCAAATCAATGATTTCTGACGCAAAATGGACCAACCTGCTTAAAGAAAAAG 3810
Db 3781 CAAACCAAATCAATGATTTCTGACGCAAAATGGACCAACCTGCTTAAAGAAAAAG - - 3838
QY 3811 GGTTTATTTAGTCAGGAAAGAGGACCTGCTTTACCCACACAGGTTCCCTCTGAGTAC 3870
Db 3839 -----AGTTTCCTCTGAGTAC 3855
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGACAGCTAGAGAAAGCC 3930
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGACAGCTAGAGAAAGCC 3915
QY 3931 CTTTCAAGAGACCGCATCGAGCTCGGTTCCGCCCGGAGGAGAGCTGCCACCGCAAGCA 3990
Db 3916 CTTTCAAGAGACCGCATCGAGCTCGGTTCCGCCCGGAGGAGAGCTGCCACCGCAAGCA 3975
QY 3991 ACGGACCAACCAACCCATCCACGCGCAGCCACCGGAGGACAGCATGCCATGTCCGCC 4050
Db 3976 ACGGACCAACCAACCCATCCACGCGCAGCCACCGGAGGACAGCATGCCATGTCCGCC 4035
QY 4051 ATCGTGGCTGCCAGAGCACAGCGCCAGTGCCATGAGCTGTGCGCCCGCCATCCAGC 4110
Db 4036 ATCGTGGCTGCCAGAGCACAGCGCCAGTGCCATGAGCTGTGCGCCCGCCATCCAGC 4095
QY 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCCGGCGTCTTAAAGGAAAGCATGCAC 4170
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCCGGCGTCTTAAAGGAAAGCATGCAC 4155
QY 4171 CACAATATTCCTCAACCGATTCACGTAAGACTGAACATGCGAGGCCAAGAGTGTGCTGTG 4230
Db 4156 CACAATATTCCTCAACCGATTCACGTAAGACTGAACATGCGAGGCCAAGAGTGTGCTGTG 4215
QY 4231 TGCTGGATACCGTGACCTTTGGAGCCAGGCATCCAAATGCTCGAATGTGACGGTGATG 4290
Db 4216 TGCTGGATACCGTGACCTTTGGAGCCAGGCATCCAAATGCTCGAATGTGACGGTGATG 4275
QY 4291 TGTCACCCCAAGTGTCCACGTGCTTTGACGCGCACCTGCGGCTTGCCCTGCTGAATATGCC 4350
Db 4276 TGTCACCCCAAGTGTCCACGTGCTTTGACGCGCACCTGCGGCTTGCCCTGCTGAATATGCC 4335
QY 4351 ACACATTCACCGAGGCCTTCTGCGGTGACAAAATGAATGCCCHAGGTCTCCAGCCAAAG 4410
Db 4336 ACACATTCACCGAGGCCTTCTGCGGTGACAAAATGAATGCCCHAGGTCTCCAGCCAAAG 4395
QY 4411 GAGCCCGACGAGCTTGCACTCGAAGGAGTGAAGTGGCCAGGAAATGAACAAAGCA 4470
Db 4396 GAGCCCGACGAGCTTGCACTCGAAGGAGTGAAGTGGCCAGGAAATGAACAAAGCA 4455
QY 4471 GGACAGCAAGGTGGGACAGGAAGTACATTTGCTGGAGGAGTCAAAAGTCTCATTTAT 4530
Db 4456 GGACAGCAAGGTGGGACAGGAAGTACATTTGCTGGAGGAGTCAAAAGTCTCATTTAT 4515

QY 4531 GACAATGAAGCCAGAGAAAGCTGGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4590
Db 4516 GACAATGAAGCCAGAGAAAGCTGGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4575
QY 4591 GACGGGATGTATCTATTTCAATGCTGCGTTGCTTCCGAACTCCGAAATACAGCCAAA 4650
Db 4576 GACGGGATGTATCTATTTCAATGCTGCGTTGCTTCCGAACTCCGAAATACAGCCAAA 4635
QY 4651 GCA----- 4653
Db 4636 GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACCACCTGCTGCGCCCGGG 4695
QY 4654 ----- 4653
Db 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4755
QY 4654 -----GAAAAAGCAGAAAGCTGATGCTTAAACTG 4680
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAGAAACAGAGCTGATGCTTAACTG 4815
QY 4681 CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4740
Db 4816 CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4875
QY 4741 CCCTTCAGTCACAGGTGTGTTGGGGCACCGAGGAAGGGCTCTACGCCCTGAATGTC 4800
Db 4876 CCCTTCAGTCACAGGTGTGTTGGGGCACCGAGGAAGGGCTCTACGCCCTGAATGTC 4935
QY 4801 TTGAAAACTCCCTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
Db 4936 TTGAAAACTCCCTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995
QY 4861 AAGCAGCTGAGAGAGCTACTCATGATAGCAGGAGAGGGGCACTGTCTCTGTGGAC 4920
Db 4996 AAGCAGCTGAGAGAGCTACTCATGATAGCAGGAGAGGGGCACTGTCTCTGTGGAC 5055
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4980
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115
QY 4981 CCCAACATTTTGAAGCTGTCAAGGGCTGCCATTTGTTGGGCGAGGCAAGATTGAGAAC 5040
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCCATTTGTTGGGCGAGGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGCATCTGTGAGGCATGCCAGCAAAAGTGTCTATTCTCGCTACAAACGAAAC 5100
Db 5176 GGGCTCTGCATCTGTGAGGCATGCCAGCAAAAGTGTCTATTCTCGCTACAAACGAAAC 5235
QY 5101 CTCAGCAAAATATGCTGATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5160
Db 5236 CTCAGCAAAATATGCTGATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
QY 5161 TTCACCAATTTACAGTATCCCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
Db 5296 TTCACCAATTTACAGTATCCCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
QY 5221 TACACGCTCGAGGAATTCCTGGAATAAGATGACCATTTCTCTGGCACCTGCTGTGTTGCC 5280
Db 5356 TACACGCTCGAGGAATTCCTGGATAAGATGACCATTTCTCTGGCACCTGCTGTGTTGCC 5415
QY 5281 GCCTCTTCCAGACGCTCCCTGTCTCAATGCTGAGGTGAACAGCGCAGGCGAGCAGAG 5340
Db 5416 GCCTCTTCCAGACGCTCCCTGTCTCAATGCTGCGAGGTGAACAGCGCAGGCGAGCAGAG 5475
QY 5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTCGTGGATTCTTACGGAACCGTAGC 5400
Db 5476 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTCGTGGATTCTTACGGAACCGTAGC 5535
QY 5401 CGCACAGACGATCTCAAGTGGAGTCCGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5460
Db 5536 CGCACAGACGATCTCAAGTGGAGTCCGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5595
QY 5461 CTGTTTGTGACCCACTTCACTCACTCGAAGTAAATTGAGATCCAGGCAAGCTCTCTCAGCA 5520

QY	5401	CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCCTTTGGCTACAGAGAACCCCTAT	5460
Db	5536	CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCCTTTGGCTACAGAGAACCCCTAT	5595
QY	5461	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGACGCTCCTCAGCA	5520
Db	5596	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGACGCTCCTCAGCA	5655
QY	5521	GGGACCCCTGCGCGAGCGTACTTGGACATCCCGAAACCCGGCGTACTCGGGCCCTGCCATT	5580
Db	5656	GGGACCCCTGCGCGAGCGTACTTGGACATCCCGAAACCCGGCGTACTCGGGCCCTGCCATT	5715
QY	5581	TCCTCAGGAGCGATTACTTTGGCGTCTCATACAGGATAAAATTAAAGGTCATTTGTCTGC	5640
Db	5716	TCCTCAGGAGCGATTACTTTGGCGTCTCATACAGGATAAAATTAAAGGTCATTTGTCTGC	5775
QY	5641	AAGGGAACCTCGTGAAGGAGTCCGGCACTGAACACCCAGCGGCCCGTCCACCTCCCGC	5700
Db	5776	AAGGGAACCTCGTGAAGGAGTCCGGCACTGAACACCCAGCGGCCCGTCCACCTCCCGC	5835
QY	5701	AGCAGCCCAACAAAGCGAGGCCACCCAGTACACAGGACATACCAAGCGCGTGGCC	5760
Db	5836	AGCAGCCCAACAAAGCGAGGCCACCCAGTACACAGGACATACCAAGCGCGTGGCC	5895
QY	5761	TCCAGGCCAGCGCGCCGAGCCCGAGCCACCCGAGAGCCAGACACACCCACCGC	5820
Db	5896	TCCAGGCCAGCGCGCCGAGCCCGAGCCACCCGAGAGCCAGACACACCCACCGC	5955
QY	5821	TACCGGAGGGGCGGACCGAGTGGCAGGGACAAGTCTCTGGCCGCCCTCGAGCGCA	5880
Db	5956	TACCGGAGGGGCGGACCGAGTGGCAGGGACAAGTCTCTGGCCGCCCTCGAGCGCA	6015
QY	5881	GAGAAGTCCCCCGCGGATGCTCAGCACCGCGAGAGAGCGGTCCCGGGAGGCTGTTT	5940
Db	6016	GAGAAGTCCCCCGCGGATGCTCAGCACCGCGAGAGAGCGGTCCCGGGAGGCTGTTT	6075
QY	5941	GAAGACAGCAGCAGGGGCGGCTGCTCGGGAGCCGTGAGGACCCCGTGTCCCCAGGTG	6000
Db	6076	GAAGACAGCAGCAGGGGCGGCTGCTCGGGAGCCGTGAGGACCCCGTGTCCCCAGGTG	6135
QY	6001	AACAGGGGAAGGGCA	6017
Db	6136	AACAGGTGAGGAGCA	6152
RESULT 6			
AX671105			
LOCUS	AX671105	6165 bp	DNA linear PAT 27-MAR-2003
DEFINITION	Sequence 1 from Patent WO03004523.		
ACCESSION	AX671105		
VERSION	AX671105.1 GI:29329569		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Zhu, Z.		
TITLE	Regulation of human citron rho/rac-interacting kinase		
JOURNAL	Patent: WO 03004523-A 1 16-JAN-2003;		
FEATURES	Bayer Aktiengesellschaft (DE)		
source	Location/Qualifiers		
	1..6165		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	86.2%; Score 5666.2; DB 6; Length 6165;		
Best Local Similarity	95.9%; Pred. No. 0;		
Matches 5947; Conservative	0; Mismatches 13; Indels 243; Gaps 3;		

QY	19	ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAACCCATT	78
Dd	1	ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAACCCATT	60
QY	79	GCCAGCCGGGCTCCAGGCTCAATCTGTTCTCCAGGGGAAACACACCTTTATGACTCAA	138
Dd	61	GCCAGCCGGGCTCCAGGCTCAATCTGTTCTCCAGGGGAAACACACCTTTATGACTCAA	120
QY	139	CAGCAGATGCTGCTCTTTCCGAGAGGGAATATAGATGCCCTCTTTGTTCTCTTTGAA	198
Dd	121	CAGCAGATGCTCTCTTTCCGAGAGGGAATATAGATGCCCTCTTTGTTCTCTTTGAA	180
QY	199	GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACCTGAGCAACTTTGTCGGAAGTAT	258
Dd	181	GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACCTGAGCAACTTTGTCGGAAGTAT	240
QY	259	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAAGGACTTCGGAAGTCAGA	318
Dd	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAAGGACTTCGGAAGTCAGA	300
QY	319	AGTCTCTGAGTTGCTGCTCTTTCGTAAGTGCAGGTGTTAAGAGAGAAGCAACCGGG	378
Dd	301	AGTCTCTGAGTTGCTGCTCTTTCGTAAGTGCAGGTGTTAAGAGAGAAGCAACCGGG	360
QY	379	GACATCTATGCTATGAAAGTATGAAGAAGAGCTTTATTTGGCCAGGAGAGGTTTCA	438
Dd	361	GACATCTATGCTATGAAAGTATGAAGAAGAGCTTTATTTGGCCAGGAGAGGTTTCA	420
QY	439	TTTTTTGAGGAGCGGAAACATATATCTCGAGACCAAGCCGCTGGATCCCAATTA	498
Dd	421	TTTTTTGAGGAGCGGAAACATATATCTCGAGACCAAGCCGCTGGATCCCAATTA	480
QY	499	CAGTATGCTTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTTGGAGG	558
Dd	481	CAGTATGCTTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTTGGAGG	540
QY	559	GACTTGCTGCTCACTTTTGAATAGATAGAGACCAAGTATGATAAACCCTGATACAGTTT	618
Dd	541	GACTTGCTGCTCACTTTTGAATAGATAGAGACCAAGTATGATAAACCCTGATACAGTTT	600
QY	619	TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTCATCTGATGGATACGTGCTATCGA	678
Dd	601	TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTCATCTGATGGATACGTGCTATCGA	660
QY	679	GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT	738
Dd	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT	720
QY	739	GGATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTTGGGACC	798
Dd	721	GGATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTTGGGACC	780
QY	799	CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACGGGGATGGAAGGACACCTAC	858
Dd	781	CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACGGGGATGGAAGGACACCTAC	840
QY	859	GGCTGAGCTGTGATGCTGGTGAAGTGGGCTGATTCCTATGAGATGATTTATGGGAGA	918
Dd	841	GGCTGAGCTGTGATGCTGGTGAAGTGGGCTGATTCCTATGAGATGATTTATGGGAGA	900
QY	919	TCCCCCTTCGAGAGGAAACCTCTGCGCAGAACCTTCAATAACATTAATGAATTTCCAGCGG	978
Dd	901	TCCCCCTTCGAGAGGAAACCTCTGCGCAGAACCTTCAATAACATTAATGAATTTCCAGCGG	960
QY	979	TTTTTGAATTTCCAGATGACCCCAAGAGTGAAGTGAAGTCTTTCTGATCTGATTTCAAAGC	1038
Dd	961	TTTTTGAATTTCCAGATGACCCCAAGAGTGAAGTGAAGTCTTTCTGATCTGATTTCAAAGC	1020
QY	1039	TTGTTGTGGGCGGAGAGGAGTGAAGTGAAGTCTTTCTGATCTTTCTGCTCCATCTTTCTTC	1098
Dd	1021	TTGTTGTGGGCGGAGAGGAGTGAAGTGAAGTCTTTCTGATCTTTCTGCTCCATCTTTCTTC	1080
QY	1099	TCTAAATTTGACTGGAAACAACTTCGTAACCTCTCTCCCTCCCTTCGTTCCCAACCTCAAG	1158

Db TCTAAATGACTGGAACAACATTCGTAATCTCTCTCCCTTCGTTCCACCTCAAG 1140
QY TCTGACGATGACACCTCCAATTTTGATGAACCGAGAGAAGATTCGTGGGTTTCATCCTCT 1218
Db TCTGACGATGACACCTCCAATTTTGATGAACCGAGAGAAGATTCGTGGGTTTCATCCTCT 1200
QY CCGTCCAGCTGAGCCCTCAGGCTCTCGGTGAAGAACTGCGGTGTTGGGTTTTCG 1278
Db CCGTCCAGCTGAGCCCTCAGGCTCTCGGTGAAGAACTGCGGTGTTGGGTTTTCG 1260
QY TACAGCAAGGCACTGGGATCTTTGGTAGATCTGAGTCTGTTGTGCGGTCTGAGCTCC 1338
Db TACAGCAAGGCACTGGGATCTTTGGTAGATCTGAGTCTGTTGTGCGGTCTGAGCTCC 1320
QY CCTGCCAGACTAGCTCCATGGGAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1398
Db CCTGCCAGACTAGCTCCATGGGAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1380
QY TCTCAGGACAAGTGTCACAGATGGAGCAAGAAATGACCCGGTTACATCGGAGGTGTCA 1458
Db TCTCAGGACAAGTGTCACAGATGGAGCAAGAAATGACCCGGTTACATCGGAGGTGTCA 1440
QY GAGGTGAGGCTGTCTTAGTCAGAAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1518
Db GAGGTGAGGCTGTCTTAGTCAGAAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500
QY TCCCTCTCGGAGCAGGACCTTGCTACCTACATCACAGAAATGCAAGTGTAAAGCGAAGT 1578
Db TCCCTCTCGGAGCAGGACCTTGCTACCTACATCACAGAAATGCAAGTGTAAAGCGAAGT 1560
QY TTGAGCAAGCAAGGATGGAGGTGCCAGGAGTGCACAAAGCACTGAGCTTCTCCAT 1638
Db TTGAGCAAGCAAGGATGGAGGTGCCAGGAGTGCACAAAGCACTGAGCTTCTCCAT 1620
QY GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1698
Db GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1680
QY GTGGAAGAAATGAGTTGATGAATCAATGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1758
Db GTGGAAGAAATGAGTTGATGAATCAATGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740
QY CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATTCAG 1818
Db CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATTCAG 1800
QY CGGAAAGCGACAGAAATGTCAGCAATAAATGTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1878
Db CGGAAAGCGACAGAAATGTCAGCAATAAATGTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1860
QY GTGGAGAAATATCGGAAATCGGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938
Db GTGGAGAAATATCGGAAATCGGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
QY CTCGAGAGAAATCGGAGAGGCT----- 1962
Db CTCGAGAGAAATCGGAGAGGCT----- 1960
QY -----GCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGGAGAACTGGAGAAACCGAG 2010
Db AATATCCGCGAGGCAAGGAGCGAGCCGAGAGGAGCTGGAGAACTGGAGAAACCGAGAG 2040
QY GATTTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAGCGCGCCATCTCTG 2070
Db GATTTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAGCGCGCCATCTCTG 2100
QY GAGAAAGAGTAAAGAGCTTAGAGCCATGGAGCCTTAGAGAAACAGACTGAAGGATGAC 2130
Db GAGAAAGAGTAAAGAGCTTAGAGCCATGGAGCCTTAGAGAAACAGACTGAAGGATGAC 2160
QY ATCCAGACAAATCCCAACAGATCCAGAGAGGCTGATTAATTTCTGGAGCTCCAGAG 2190

Db ATCCAGACAAATCCCAACAGATCCAGAGAGGCTGATATAAATTTCTGGAGCTCGAAGAG 2220
QY AAACATCGGAGGCCCAAGTCTCAGCCAGAGCACTAGAAAGTGCACTTGAAACAGAAAGAG 2250
Db AAACATCGGAGGCCCAAGTCTCAGCCAGAGCACTAGAAAGTGCACTTGAAACAGAAAGAG 2280
QY CAGCACTATGAGGAAAGATTAAGTGTGGAACAATCAGATAAAAGAAAGACCTGGCTGAC 2310
Db CAGCACTATGAGGAAAGATTAAGTGTGGAACAATCAGATAAAAGAAAGACCTGGCTGAC 2340
QY AAGGAGACACTGGAGAAACATGATCAGAGACA CAGAGGAGGAGGCCATGAGAAAGGGCAAA 2370
Db AAGGAGACACTGGAGAAACATGATCAGAGACA CAGAGGAGGAGGCCATGAGAAAGGGCAAA 2400
QY ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGAA 2430
Db ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGAA 2460
QY CAGAGGATTTGGAACTGTCTGAAAGCAATAAATCTTCAGCAATAGCAGTCTTTTACC 2490
Db CAGAGGATTTGGAACTGTCTGAAAGCAATAAATCTTCAGCAATAGCAGTCTTTTACC 2520
QY CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550
Db CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
QY CTGGAGACACAGGCTGGGAAAGTTGGAGGCCAGAAACCAAACTGGAGGAGCAGCTGGAG 2610
Db CTGGAGACACAGGCTGGGAAAGTTGGAGGCCAGAAACCAAACTGGAGGAGCAGCTGGAG 2640
QY AAGATCAGCCACCAAGACCAAGTGAACAAGAACTGGCTGTGGAACCTGGAGACAAGATTG 2670
Db AAGATCAGCCACCAAGACCAAGTGAACAAGAACTGGCTGTGGAACCTGGAGACAAGATTG 2700
QY CGGAGGTCAGTCTAGAGCAGGAGCAGAACTGGAGCTCAAGCCAGCTCACAAGAG 2730
Db CGGAGGTCAGTCTAGAGCAGGAGCAGAACTGGAGCTCAAGCCAGCTCACAAGAG 2760
QY CTAAGTCTCTCTGAGGAGCCGAGTCAAGTTGACAGCTTGCAGGCTGCAACGGGCG 2790
Db CTAAGTCTCTCTGAGGAGCCGAGTCAAGTTGACAGCTTGCAGGCTGCAACGGGCG 2820
QY GCCCTGAGAGCCAGCTTCCGAGGCAAGACAGAGCTGGAAGACCAAGAGAGCT 2850
Db GCCCTGAGAGCCAGCTTCCGAGGCAAGACAGAGCTGGAAGACCAAGAGAGCT 2880
QY GAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGGCAATTTGATGCT 2910
Db GAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGGCAATTTGATGCT 2940
QY CTTGCTAACAGCTGTACTGTAACTCAGAGCTGGAGAGCAGCTTAACAGCTGACCGAG 2970
Db CTTGCTAACAGCTGTACTGTAACTCAGAGCTGGAGAGCAGCTTAACAGCTGACCGAG 3000
QY GACAACCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGTTCT 3030
Db GACAACCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGTTCT 3060
QY GGCCCAACAGAGATTTGACAACTGCGAAGTGAAGTGAACCATCTCCCGCGGAGATC 3090
Db GGCCCAACAGAGATTTGACAACTGCGAAGTGAAGTGAACCATCTCCCGCGGAGATC 3120
QY ACGGAAACAGAGATGAGCTTACAGCCAGAAACCAACAGATGAGGCTCTGAAGACACG 3150
Db ACGGAAACAGAGATGAGCTTACAGCCAGAAACCAACAGATGAGGCTCTGAAGACACG 3180
QY TGCAACATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210
Db TGCAACATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240
QY AAAGAGCGGAGTGGAGGCGCTCGAGAGGCTCTCGGTTGATGAGAAATCCCAAGTTTGG 3270
Db AAAGAGCGGAGTGGAGGCGCTCGAGAGGCTCTCGGTTGATGAGAAATCCCAAGTTTGG 3300

QY 3271 TGTCCGGTTCCAGAGCTGCAGAGATGCTGGACACCGAGAAACAGAGAGGCGGAGAGCC 3330
DB 3301 TGTCCGGTTCCAGAGCTGCAGAGATGCTGGACACCGAGAAACAGAGAGGCGGAGAGCC 3360
QY 3331 GATCAGCGGATCACCAGAGTCTCCAGAGTGTGGAGCTGGAGCTGGAGAGCACAAGGCT 3390
DB 3361 GATCAGCGGATCACCAGAGTCTCCAGAGTGTGGAGCTGGAGAGCACAAGGCT 3420
QY 3391 GAGATTCTCGCTCTGCAGAGGCTCTCAAGAGCAGAGCTGAAGGCGGAGAGCTCTCT 3450
DB 3421 GAGATTCTCGCTCTGCAGAGGCTCTCAAGAGCAGAGCTGAAGGCGGAGAGCTCTCT 3480
QY 3451 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGAATGCCGAGAGCTTA 3510
DB 3481 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGAATGCCGAGAGCTTA 3540
QY 3511 CAGCAGAGCTGGAGAGTGAACAGAGCTCAACAGAGAGCTTCTGAGAGAGCAAGCCAAA 3570
DB 3541 CAGCAGAGCTGGAGAGTGAACAGAGCTCAACAGAGAGCTTCTGAGAGAGCAAGCCAAA 3600
QY 3571 TTACAGCAGCAGATGAGCTGCAGAGAAATACATTTTCGCTGACTCAAGGACTGCAA 3630
DB 3601 TTACAGCAGCAGATGAGCTGCAGAGAAATACATTTTCGCTGACTCAAGGACTGCAA 3660
QY 3631 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTATCAGCTG 3690
DB 3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTATCAGCTG 3720
QY 3691 GAAAAATTCAGGTTCTCTATTCTCATGAAAAAGGTGAAAAAGGAGGAGCTATTTCTCAA 3750
DB 3721 GAAAAATTCAGGTTCTCTATTCTCATGAAAAAGGTGAAAAAGGAGGAGCTATTTCTCAA 3780
QY 3751 CAAACAAACTCAATGATTTCTGAGCAACAAATGAGCAACCTGCTTAAAAAGAAAAAG 3810
DB 3781 CAAACAAACTCAATGATTTCTGAGCAACAAATGAGCAACCTGCTTAAAAAGAAAA-- 3838
QY 3811 GGTATTATTAGTCGAGGAAAGAGGACCTGCTTTACCCACACAGGTTCTCTGAGTAC 3870
DB 3839 -----AGTTTCTCTGAGTAC 3855
QY 3871 AATGAGCTGAAGTGCCTGGAGAGGAGAAAGTCTGCTGTCAGAGCTAGAGAAAGCC 3930
DB 3856 AATGAGCTGAAGTGCCTGGAGAGGAGAAAGTCTGCTGTCAGAGCTAGAGAAAGCC 3915
QY 3931 CTTCAGAGAACCGGATCAGCTCCGGTCCGCGCGGAGAGAGCTGCCACCGCAAGCA 3990
DB 3916 CTTCAGAGAACCGGATCAGCTCCGGTCCGCGCGGAGAGAGCTGCCACCGCAAGCA 3975
QY 3991 ACGGACCAACCAACCCATCCACGCCAGCAACCGGAGGAGCAGATCGCCATGTCGCGC 4050
DB 3976 ACGGACCAACCAACCCATCCACGCCAGCAACCGGAGGAGCAGATCGCCATGTCGCGC 4035
QY 4051 ATCGTGGGTGCGCAGAGCAACAGCCAGTGCATGAGCTGCTGGCCCGCCGATCCAGC 4110
DB 4036 ATCGTGGGTGCGCAGAGCAACAGCCAGTGCATGAGCTGCTGGCCCGCCGATCCAGC 4095
QY 4111 CGCAGAAAGGAGTCTCACTCCAGAGGAATTTAGTGGCGCTCTTAAGGAAACGCAATGCAC 4170
DB 4096 CGCAGAAAGGAGTCTCACTCCAGAGGAATTTAGTGGCGCTCTTAAGGAAACGCAATGCAC 4155
QY 4171 CACAAATTTCTCACCAGATTCACGTAAGGACTGAACATCGAGCCCAAAAGTGTGCTGTG 4230
DB 4156 CACAAATTTCTCACCAGATTCACGTAAGGACTGAACATCGAGCCCAAAAGTGTGCTGTG 4215
QY 4231 TGTCTGGATACCGTGACATTTGAGCCGAGGATCCAAATGCTCGAATGTCAGGATGATG 4290
DB 4216 TGTCTGGATACCGTGACATTTGAGCCGAGGATCCAAATGCTCGAATGTCAGGATGATG 4275
QY 4291 TGTCAACCCCAAGTGTCCAGTCTTGGCCAGCCACCTGGGGCTTGGCTGCTGAATATGCC 4350
DB 4276 TGTCAACCCCAAGTGTCCAGTCTTGGCCAGCCACCTGGGGCTTGGCTGCTGAATATGCC 4335

QY 4351 ACACACTTCCAGAGGCTTCTGCGGTGACAAATAGAACTCCCGAGTCTCCAGACCAAG 4410
DB 4336 ACACACTTCCAGAGGCTTCTGCGGTGACAAATAGAACTCCCGAGTCTCCAGACCAAG 4395
QY 4411 GAGCCAGCAGCAGCTTGCACCTGGAAGGTGATGAAGTGCAGGAATAACAAACGA 4470
DB 4396 GAGCCAGCAGCAGCTTGCACCTGGAAGGTGATGAAGTGCAGGAATAACAAACGA 4455
QY 4471 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTTGAGGGATCAAAAGTCTCATTTAT 4530
DB 4456 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTTGAGGGATCAAAAGTCTCATTTAT 4515
QY 4531 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4590
DB 4516 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4575
QY 4591 GACGGGATGATATCTATTCTATGCTGCGTTGGTCTTCCGAACTCGCAATACAGCCAAA 4650
DB 4576 GACGGGATGATATCTATTCTATGCTGCGTTGGTCTTCCGAACTCGCAATACAGCCAAA 4635
QY 4651 GCA----- 4653
DB 4636 GCAGATGCTCCATACATACTGAAGAGGAATCTCAOCCGACACACCCTGCTGGCCCGGG 4695
QY 4654 ----- 4653
DB 4696 AGAACCTTACTTGTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACCGCTTA 4755
QY 4654 -----GAAAAAGCAGAGCTGATCTAAACTG 4680
DB 4756 GAATCAGTGTGCGCAGGTGGAGAGTTTCTAGGGAAAAAGCAGAAAGTGTGCTAAACTG 4815
QY 4681 CTTGGAAGCTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACTGCAAGCTG 4740
DB 4816 CTTGGAAGCTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACTGCAAGCTG 4875
QY 4741 CCTTTCAGTGACCAAGTGTGTGTTGTCGGCAACGAGGAAGGGTCTACGCGCTGAATGTC 4800
DB 4876 CCTTTCAGTGACCAAGTGTGTGTTGTCGGCAACGAGGAAGGGTCTACGCGCTGAATGTC 4935
QY 4801 TTGAAAACTCCCTTAAACCATGTCAGGAAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
DB 4936 TTGAAAACTCCCTTAAACCATGTCAGGAAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995
QY 4861 AAGGACCTGGAGAAAGTACTCATGATAGCAGGAGAGAGGCGGACATGTGTCTTGTGAC 4920
DB 4996 AAGGACCTGGAGAAAGTACTCATGATAGCAGGAGAGAGGCGGACATGTGTCTTGTGAC 5055
QY 4921 GTGAGAAAGTGAACAGTCTCGGCCAGTCCCACTGCTGCTGCCAGCCGACATCTCA 4980
DB 5056 GTGAGAAAGTGAACAGTCTCGGCCAGTCCCACTGCTGCTGCCAGCCGACATCTCA 5115
QY 4981 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5040
DB 5116 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTCTGCTATTTCTCGCTACAGCAAAAC 5100
DB 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTCTGCTATTTCTCGCTACAGCAAAAC 5235
QY 5101 CTCAGCAAACTCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160
DB 5236 CTCAGCAAACTCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
QY 5161 TTCACCAATTAAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
DB 5296 TTCACCAATTAAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
QY 5221 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCATTCCTTGGCACCTGCTGTGTTGCC 5280
DB 5356 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCATTCCTTGGCACCTGCTGTGTTGCC 5415
QY 5281 GCCTCTTCCACAGCTTCCCTGTCTCAATCTGTCAGGTGAACAGCGGAGCGGAGAG 5340

Db	5416	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGGTGAACACGCGAGGGCAGCGAGAG	5475
Qy	5341	GAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGACGTAGC	5400
Db	5476	GAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGACGTAGC	5535
Qy	5401	CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCATTGCGCTACAGAAACCCCTAT	5460
Db	5536	CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCATTGCGCTACAGAAACCCCTAT	5595
Qy	5461	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACACGCTCTCAGCA	5520
Db	5596	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACACGCTCTCAGCA	5655
Qy	5521	GGGACCCCTGCGAGAGGTAACCTGGACATCCGAAACCGGCGCTACCTGGGCCCTGCCATT	5580
Db	5656	GGGACCCCTGCGAGAGGTAACCTGGACATCCGAAACCGGCGCTACCTGGGCCCTGCCATT	5715
Qy	5581	TCCTCAGGAGGGATTACTTTGGCGTCTCATACGAGATAAATTAAGGTTCATTTGCTGC	5640
Db	5716	TCCTCAGGAGGGATTACTTTGGCGTCTCATACGAGATAAATTAAGGTTCATTTGCTGC	5775
Qy	5641	AAGGGAACCTCGTGAAGAGTCGGGCACTGAACACCAACCGGGCGCCGTCACACTCCCGC	5700
Db	5776	AAGGGAACCTCGTGAAGAGTCGGGCACTGAACACCAACCGGGCGCCGTCACACTCCCGC	5835
Qy	5701	AGCAGCCCCAACAGCGAGGCCACCCACGTAACAAGAGCACATCAACAGCGCGTGGCC	5760
Db	5836	AGCAGCCCCAACAGCGAGGCCACCCACGTAACAAGAGCACATCAACAGCGCGTGGCC	5895
Qy	5761	TCCAGCCCCAGCGCGCCGAAGGCCCCAGCCACCCGCGAGAGCCAGCAACCCCAACCCG	5820
Db	5896	TCCAGCCCCAGCGCGCCGAAGGCCCCAGCCACCCGCGAGAGCCAGCAACCCCAACCCG	5955
Qy	5821	TACCGCGAGGGCGGACCGAGCTGCGCAGGGACAAGTCTCTGCGCGCCGCCCTTGGAGCGA	5880
Db	5956	TACCGCGAGGGCGGACCGAGCTGCGCAGGGACAAGTCTCTGCGCGCCGCCCTTGGAGCGA	6015
Qy	5881	GAGAGTCCCCCGGCCGGATGCTTCAGCACCGCGAGAGAGCGGTCTCCCGGAGGCTGTTT	5940
Db	6016	GAGAGTCCCCCGGCCGGATGCTTCAGCACCGCGAGAGAGCGGTCTCCCGGAGGCTGTTT	6075
Qy	5941	GAAGACAGCAGCAGGGGCGGGCTGCTTGGCGGAGCCGTGAGGACCCCGCTGTCCCAAGTGTG	6000
Db	6076	GAAGACAGCAGCAGGGGCGGGCTGCTTGGCGGAGCCGTGAGGACCCCGCTGTCCCAAGTGTG	6135
Qy	6001	AACAAGGAAGAGGGCAGAGTGC	6023
Db	6136	AACAAGGCTTGGACACAGTCTTC	6158

ORIGIN	Query Match Best Local Similarity Matches 5944; Conservative	86.1%; 95.8%; 0;	Score 5661.4; Pred. No. 0; Mismatches 16;	DB 6; Length 6165; Indels 243; Gaps 3;
QY	19	ATGTTGAAGTTC	AAATATGAGCGCGGAATCTTTTGATGCTGGTCTGCTGCTGAACCCATT	78
DB	1	ATGTTGAAGTTC	AAATATGAGCGCGGAATCTTTTGATGCTGGTCTGCTGCTGAACCCATT	60
QY	79	GCCAGCGCGCTCC	CAGGCTGAATCTGTTCTCCAGGGGAACCCACCCCTTATGACTCAA	138
DB	61	GCCAGCGCGCTCC	CAGGCTGAATCTGTTCTCCAGGGGAACCCACCCCTTATGACTCAA	120
QY	139	CAGCAGATGTC	TCTCCTCTTTCCGAGAGGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA	198
DB	121	CAGCAGATGTC	TCTCCTCTTTCCGAGAGGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA	180
QY	199	GAATGCAGTCAG	CGCTGCTGTGATGAAGATTAAGCACGTGAGCAACTTGTTCGGGAAGTAT	258
DB	181	GAATGCAGTCAG	CGCTGCTGTGATGAAGATTAAGCACGTGAGCAACTTGTTCGGGAAGTAT	240
QY	259	TCCGACACCATAG	CTAGTTTACAGGAGCTCCAGCCCTTCGGCAAAAGCACTTCGAAGTCAGA	318
DB	241	TCCGACACCATAG	CTAGTTTACAGGAGCTCCAGCCCTTCGGCAAAAGCACTTCGAAGTCAGA	300
QY	319	AGTCTTGTAGG	TGTGGTCACTTTTGCTGAAGTCAGGTGGTAAAGAGAAAGCAACCCGG	378
DB	301	AGTCTTGTAGG	TGTGGTCACTTTTGCTGAAGTCAGGTGGTAAAGAGAAAGCAACCCGG	360
QY	379	GACATCTATCT	ATGAAAGTGATGAAGAAGAGGCTTTATTCGCCCAGAGCAGGTTTCA	438
DB	361	GACATCTATCT	ATGAAAGTGATGAAGAAGAGGCTTTATTCGCCCAGAGCAGGTTTCA	420
QY	439	TTTTTTGAGGA	AGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA	498
DB	421	TTTTTTGAGGA	AGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA	480
QY	499	CAGTATGCC	TTTCAGACAAAATCACCTTTATCTGATCGRGGGAATATCAGCGTCGGGG	558
DB	481	CAGTATGCC	TTTCAGACAAAATCACCTTTATCTGATCGRGGGAATATCAGCGTCGGGG	540
QY	559	GACTTCCTG	CTCATTTTGAATAGATATGAGGACCAAGTTAGATGAAAACTGATACAGTTT	618
DB	541	GACTTCCTG	CTCATTTTGAATAGATATGAGGACCAAGTTAGATGAAAACTGATACAGTTT	600
QY	619	TACCTAGCTAG	GTGATTTTGGCTGTTTCACAGCGTTTCATCTGATGGGATACGTGATCGA	678
DB	601	TACCTAGCTAG	GTGATTTTGGCTGTTTCACAGCGTTTCATCTGATGGGATACGTGATCGA	660
QY	679	GACATCAAG	CCCTGAGAACTTCTCGTTACCCACAGGACACATCAAGCTGGTGGATTTT	738
DB	661	GACATCAAG	CCCTGAGAACTTCTCGTTACCCACAGGACACATCAAGCTGGTGGATTTT	720
QY	739	GGATCTGCC	CGGAAATGAAATTCAAACAAGATGGTGAATGCCAAACTCCCGATGGGACC	798
DB	721	GGATCTGCC	CGGAAATGAAATTCAAACAAGATGGTGAATGCCAAACTCCCGATGGGACC	780
QY	799	CCAGATTAC	TGATGGCTTCAAGTGTGATGAACGGGGATGGAAGGACCACTAC	858
DB	781	CCAGATTAC	TGATGGCTTCAAGTGTGATGAACGGGGATGGAAGGACCACTAC	840
QY	859	GGCTTGA	CTGTGACTGTGGCTGATGGCGGTGATGCTATGAGATGATTTATGGGAGA	918
DB	841	GGCTTGA	CTGTGACTGTGGCTGATGGCGGTGATGCTATGAGATGATTTATGGGAGA	900
QY	919	TCCCCCTTC	CGCAGAGGGAACCTCTCCAGAACCTTCAATTAATTTATCAATTTCCAGCGG	978
DB	901	TCCCCCTTC	CGCAGAGGGAACCTCTCCAGAACCTTCAATTAATTTATCAATTTCCAGCGG	960
QY	979	TTTTTGA	ATTTCCAGATGACCCCCAAGTGAGCAGTGACTTTCTTTGATCTGATTCAAAGC	1038
DB	961	TTTTTGA	ATTTCCAGATGACCCCCAAGTGAGCAGTGACTTTCTTTGATCTGATTCAAAGC	1020

QY 1039 TTGTTGTGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTGTCATCTCTTTCTTC 1098
Db 1021 TTGTTGTGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTGTCATCTCTTTCTTC 1080
QY 1099 TCTAAAATTGAGTGGAAACAACTTTCGTAACCTCTCTCCCTCCCTTCGTTCCACCTCAAG 1158
Db 1081 TCTAAAATTGAGTGGAAACAACTTTCGTAACCTCTCTCCCTCCCTTCGTTCCACCTCAAG 1140
QY 1159 TCTGACGATGACACCTCCCAATTTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCTCT 1218
Db 1141 TCTGACGATGACACCTCCCAATTTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCTCT 1200
QY 1219 CCGTGCCACTGAGCCCTCAGGCTTCGCGGTGAAGAACTCCCGTTTGTGGGTTTTCG 1278
Db 1201 CCGTGCCACTGAGCCCTCAGGCTTCGCGGTGAAGAACTCCCGTTTGTGGGTTTTCG 1260
QY 1279 TACAGCAAGCACTGGGATTCCTGTAGATCTGAGTCTGTTGTGCGGTCTGACTCC 1338
Db 1261 TACAGCAAGCACTGGGATTCCTGTAGATCTGAGTCTGTTGTGCGGTCTGACTCC 1320
QY 1339 CTTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1398
Db 1321 CTTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
QY 1399 TCTCAGCAAGTGTCAAGATGAGCAGGAAATGACCCGTTTACATCGGAGGTGCA 1458
Db 1381 TCTCAGCAAGTGTCAAGATGAGCAGGAAATGACCCGTTTACATCGGAGGTGCA 1440
QY 1459 GAGTTCGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTCAGACTCAGAGA 1518
Db 1441 GAGTTCGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTCAGACTCAGAGA 1500
QY 1519 TCCCTCTCGAGCAGACCTTGCTACTACATCAAGATGAGTGTAAAGGGAAGT 1578
Db 1501 TCCCTCTCGAGCAGACCTTGCTACTACATCAAGATGAGTGTAAAGGGAAGT 1560
QY 1579 TTGGACACAGCAGGATGAGGTGTCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT 1638
Db 1561 TTGGACACAGCAGGATGAGGTGTCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT 1620
QY 1639 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAAGAGCAGAGTACAGGCTCAA 1698
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAAGAGCAGAGTACAGGCTCAA 1680
QY 1699 GTGGAGAAATGAGTGTAGTGAATCAGTTGGAGAGATCTCTCAGCAGAAGA 1758
Db 1681 GTGGAGAAATGAGTGTAGTGAATCAGTTGGAGAGATCTCTCAGCAGAAGA 1740
QY 1759 CGGAGTGAATCTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGTAAGAAATCAAAG 1818
Db 1741 CGGAGTGAATCTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGTAAGAAATCAAAG 1800
QY 1819 CGGAAGCGACAGAAATGTCAGATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1878
Db 1801 CGGAAGCGACAGAAATGTCAGATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1860
QY 1879 GTGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938
Db 1861 GTGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
QY 1939 CTCAAGAGAACTGGAGAGGCT----- 1962
Db 1921 CTCAAGAGAACTGGAGAGGCTGTAAGAACCCAGCAGCAGGAGCCACCGAGCTGTCGAG 1980
QY 1963 -----GCARAAGGAGCGAGCGAGAGGAGCTGAGAGCTGCAGAACCGAGAG 2010
Db 1981 AATATCCCGAGGCAAGAGGAGCGAGCGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2040
QY 2011 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTCAGGAAACCGCCCAATCTCTG 2070
Db 2041 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTCAGGAAACCGCCCAATCTCTG 2100

QY 2071 GAGAACAAAGGTAAAGAGACTAGAGACCATGAGCGGTAGAGAAAACAGACTGAAGGTGAC 2130
Db 2101 GAGAACAAAGGTAAAGAGACTAGAGACCATGAGCGGTAGAGAAAACAGACTGAAGGTGAC 2160
QY 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2190
Db 2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2220
QY 2191 AAACATCGGAGGCCCCAAGTCTCAGCCAGCACTTAGAAGTGCACCTGAAACAGAAAGAG 2250
Db 2221 AAACATCGGAGGCCCCAAGTCTCAGCCAGCACTTAGAAGTGCACCTGAAACAGAAAGAG 2280
QY 2251 CAGCACTATCAGGAAAAGATTAAAGTCTTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2310
Db 2281 CAGCACTATCAGGAAAAGATTAAAGTCTTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340
QY 2311 AAGGAGACACTGGAGAACATGATGACAGAGACACGAGAGGAGGCCCATGAGAGGGCAAA 2370
Db 2341 AAGGAGACACTGGAGAACATGATGACAGAGACACGAGAGGAGGCCCATGAGAGGGCAAA 2400
QY 2371 ATTTCTCAGCNAACAGAGGCGATGATCAATGCTATGATTCCAGATCAGATCCCTCGAA 2430
Db 2401 ATTTCTCAGCNAACAGAGGCGATGATCAATGCTATGATTCCAGATCAGATCCCTCGAA 2460
QY 2431 CAGAGGATTTGTGGAATCTGTGAGGCCAATAAATTTGCAGCAATAGCAGTCTTTTACC 2490
Db 2461 CAGAGGATTTGTGGAATCTGTGAGGCCAATAAATTTGCAGCAATAGCAGTCTTTTACC 2520
QY 2491 CAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGCNAACAGAAATTTTAC 2550
Db 2521 CAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGCNAACAGAAATTTTAC 2580
QY 2551 CTGAGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG 2610
Db 2581 CTGAGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG 2640
QY 2611 AAGATCAGCCACCAAGACCAAGTGAAGAATCGGCTGTGGAATCTGGAGCAAGATTG 2670
Db 2641 AAGATCAGCCACCAAGACCAAGTGAAGAATCGGCTGTGGAATCTGGAGCAAGATTG 2700
QY 2671 CGGAGGCTCAGCTTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAGAG 2730
Db 2701 CGGAGGCTCAGCTTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAGAG 2760
QY 2731 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGAAGCCCTGCAAGGCTGCAAGGCG 2790
Db 2761 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGAAGCCCTGCAAGGCTGCAAGGCG 2820
QY 2791 GCGCTGGAGCAGCTTCGCCAGCGAAGCAGAGCTGGAGAGACCAAGCAGAGAGCT 2850
Db 2821 GCGCTGGAGCAGCTTCGCCAGCGAAGCAGAGCTGGAGAGACCAAGCAGAGAGCT 2880
QY 2851 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2910
Db 2881 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940
QY 2911 CTTTGTAAACAGCTGTATGTAAATCAGACCTGGAGAGAGCTAAACAGCTCAGCGAG 2970
Db 2941 CTTTGTAAACAGCTGTATGTAAATCAGACCTGGAGAGAGCTAAACAGCTCAGCGAG 3000
QY 2971 GACACGCTGAACCTCAACCAAACTTTCTACTTTGTCAAAACAACTCGATGAGGCTTCT 3030
Db 3001 GACACGCTGAACCTCAACCAAACTTTCTACTTTGTCAAAACAACTCGATGAGGCTTCT 3060
QY 3031 GCGCCAAACAGAGATTGTACAACTGCGAAAGTGAAGTGGACCAATCTCCCGCGGAGATC 3090
Db 3061 GCGCCAAACAGAGATTGTACAACTGCGAAAGTGAAGTGGACCAATCTCCCGCGGAGATC 3120
QY 3091 ACGGAACAGAGATGACAGCTTACAGCCAGAGCAACAGATGGAGGCTCTCAGAGCAGC 3150
Db 3121 ACGGAACAGAGATGACAGCTTACAGCCAGAGCAACAGATGGAGGCTCTCTGAAGCAGC 3180
QY 3151 TGCACCATGCTGGAGGAACAGGTCATGATTGGAGGCCCTAAAACGATGAGCTGCTAGAA 3210

Db	3181		TCACATGCTGGAGAAACAGTCA	TTGGAGGCCCTAAACGATGAGCTGCTAGAA	3240
QY	3211		AAAGAGCGCAGTGGAGGCCCTGGAGGACGCTCTCGGTGATGAGAAATCCAGTTTGAG		3270
Db	3241		AAAGAGCGCAGTGGAGGCCCTGGAGGACGCTCTGGGTGATGAGAAATCCAGTTTGAG		3300
QY	3271		TGTCGGGTTCGAGACTCGAGAGATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCC		3330
Db	3301		TGTCGGGTTCGAGACTCGAGAGATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCC		3360
QY	3331		GATCAGCGGATCACCGAGCTCTCGCCAGCTGGAGCTGGCAGTGAAGGAGCACAAAGCT		3390
Db	3361		GATCAGCGGATCACCGAGCTCTCGCCAGCTGGAGCTGGCAGTGAAGGAGCACAAAGCT		3420
QY	3391		GAGATTCTCGTCTGACGACGCTCTCAAAGAGCAGAAGCTGAAGGCCGAGAGCCTCTCT		3450
Db	3421		GAGATTCTCGTCTGACGACGCTCTCAAAGAGCAGAAGCTGAAGGCCGAGAGCCTCTCT		3480
QY	3451		GACAGCTCAATGATCGGACGAAGCATGCTATGCTTGAATGATGATGCCGAAGCTTA		3510
Db	3481		GACAGCTCAATGATCGGACGAAGCATGCTATGCTTGAATGATGATGCCGAAGCTTA		3540
QY	3511		CAGCAGAAGCTGGAGACTGAACGAGACTCAAACAGAGGCTTCGGAAGAGCAAGCCAA		3570
Db	3541		CAGCAGAAGCTGGAGACTGAACGAGACTCAAACAGAGGCTTCGGAAGAGCAAGCCAA		3600
QY	3571		TTACAGCAGCAGATGGACCTGCAGAAAATTCACATTTTCGCTCTGACTCAAGACTGCAG		3630
Db	3601		TTACAGCAGCAGATGGACCTGCAGAAAATTCACATTTTCGCTCTGACTCAAGACTGCAG		3660
QY	3631		GAAGCTCTAGATCGGGTGATCTACTCAAGACAGAAAAGTGAATGGAGTATCAGCTG		3690
Db	3661		GAAGCTCTAGATCGGGTGATCTACTCAAGACAGAAAAGTGAATGGAGTATCAGCTG		3720
QY	3691		GAAGCATTAGGTTCTCTATTCTCATGAAAAGTGAAATGGAGACGACTATTTCTCAA		3750
Db	3721		GAAGCATTAGGTTCTCTATTCTCATGAAAAGTGAAATGGAGACGACTATTTCTCAA		3780
QY	3751		CAAAACCAACTCATGTATTTCTGCAAGCAAAATGACCAACTGCTAAAAAGAAAAG		3810
Db	3781		CAAAACCAACTCATGTATTTCTGCAAGCAAAATGACCAACTGCTAAAAAGAAAAG		3838
QY	3811		GGTTTATTAGTCGACGGAAAGAGACCCCTGCTTTACCCACACAGTTCCTCTGCAGTAC		3870
Db	3839		-----AGGTTCTCTGCAGTAC		3855
QY	3871		AATGAGCTGAAGCTGGCCCTGGAGAGGAAAGCTCGCTGTCAGAGCTAGAGNAGCC		3930
Db	3886		AATGAGCTGAAGCTGGCCCTGGAGAGGAAAGCTCGCTGTCAGAGCTAGAGNAGCC		3915
QY	3931		CTTCAGAAAGCCCGCATCGAGCTCCGTCGCGCCGGAGGAAGCTCCCAACCGCAAGCA		3990
Db	3916		CTTCAGAAAGCCCGCATCGAGCTCCGTCGCGCCGGAGGAAGCTCCCAACCGCAAGCA		3975
QY	3991		ACGACCAACCCACACCATCCAGCCAGCCACCGCAGGAGCAGATCGCATGTCGCGC		4050
Db	3976		ACGACCAACCCACACCATCCAGCCAGCCACCGCAGGAGCAGATCGCATGTCGCGC		4035
QY	4051		ATGTCGGTTCGCCAGAGCACACCGCCAGTGCCATGAGCTGCTGCGCCCGCCATCCAGC		4110
Db	4036		ATGTCGGTTCGCCAGAGCACACCGCCAGTGCCATGAGCTGCTGCGCCCGCCATCCAGC		4095
QY	4111		CGCAGAAAGAGTCTTCAACTCAGAGGAATTTAGTCGGGTCTTTAAGGAACGATGAC		4170
Db	4096		CGCAGAAAGAGTCTTCAACTCAGAGGAATTTAGTCGGGTCTTTAAGGAACGATGAC		4155
QY	4171		CACAAATTCTCACCGATTCAACGTAGGACTGAACATCGAGGCCACAAAGTGTCTGTG		4230
Db	4156		CACAAATTCTCACCGATTCAACGTAGGACTGAACATCGAGGCCACAAAGTGTCTGTG		4215
QY	4231		TGTTCTGATACCGTGACCTTTGACCGCCAGGCATCCAAATGCTTCGAATGTCAGTGTATG		4290

Db	4216	TGTCTGGATACCGTGCACCTTTGGACGCCACGGCATCCAAATGTCTCGAATGTCTCAGGTGATG	4276
Qy	4291	TGTCAACCCCAAGTCTCCAGTGTCTTGCCAGCCACTTGCGGCTTGCTGCTGCTGAATATGCC	4350
Db	4276	TGTCAACCCCAAGTCTCCAGTGTCTTGCCAGCCACTTGCGGCTTGCTGCTGCTGAATATGCC	4335
Qy	4351	ACACACTTCCCGAGAGGCTTCTCGGTGACAAATGAAATCTCCCAAGTTCAGACCAAG	4410
Db	4336	ACACACTTCCCGAGAGGCTTCTCGGTGACAAATGAAATCTCCCAAGTTCAGACCAAG	4395
Qy	4411	GAGCCACGACGACGCTTTCACCTTGGAAAGGTGATGAAGTGTCCCAAGGAATACAAACGA	4470
Db	4396	GAGCCACGACGACGCTTTCACCTTGGAAAGGTGATGAAGTGTCCCAAGGAATACAAACGA	4455
Qy	4471	GGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGGATCAAAAGTCTCTCATTTAT	4530
Db	4456	GGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGGATCAAAAGTCTCTCATTTAT	4515
Qy	4531	GACAACTACGCCACAGAGAGCTTGACACAGGCGCGGTGGAAGATTTTTCAGCTGTGCCTTCCC	4590
Db	4516	GACAACTACGCCACAGAGAGCTTGACACAGGCGCGGTGGAAGATTTTTCAGCTGTGCCTTCCC	4575
Qy	4591	GACGGGATGTATCTATTTTCATGTGTCGGTGTGCTTCCGAACTCCCAATACAGCCAAA	4650
Db	4576	GACGGGATGTATCTATTTTCATGTGTCGGTGTGCTTCCGAACTCCCAATACAGCCAAA	4635
Qy	4651	GCA-----	4653
Db	4636	GCAGATGTCCCATACATACCTAGAGATGGAATCTCACCCGCACACACCTGCTGGCCGGG	4695
Qy	4654	-----	4653
Db	4696	AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGTGACACACACGGCTGGTCAACCGCTTA	4755
Qy	4654	-----	4680
Db	4756	GAATCAGTTGTGCGCAGGTGGGAGAGTTTCTAGGGGAAAAGCAGAGCTGATGCTAAACTG	4815
Qy	4681	CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTG	4740
Db	4816	CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTG	4875
Qy	4741	CCCTTCAGTGACCAAGTGTGTGTGGGCACCCGAGGAAGGCTCTACGCCCTGAATGTC	4800
Db	4876	CCCTTCAGTGACCAAGTGTGTGTGGGCACCCGAGGAAGGCTCTACGCCCTGAATGTC	4935
Qy	4801	TTGAAAACTCCCTTAACCATCTGCCAGGAATTGAGCAGCTTTCCTCAATTTATATTATC	4860
Db	4936	TTGAAAACTCCCTTAACCATCTGCCAGGAATTGAGCAGCTTTCCTCAATTTATATTATC	4995
Qy	4861	AAGGACCTCGGAAGCTACTCATGTAGCAGAGAAGACCGGGCACTGTCTTGTGGAC	4920
Db	4996	AAGGACCTCGGAAGCTACTCATGTAGCAGAGAAGACCGGGCACTGTCTTGTGGAC	5055
Qy	4921	GTGAGAAAGTGAACAGTCCCTGCCCGACATCCACCTGCTGCCACCGCAGCATCTCA	4980
Db	5056	GTGAGAAAGTGAACAGTCCCTGCCCGACATCCACCTGCTGCCACCGCAGCATCTCA	5115
Qy	4981	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTGAGAAC	5040
Db	5116	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTGAGAAC	5175
Qy	5041	GGGCTCTGATCTGTGCGAGCATGCCACAGCAAAAGTGTGTCATTTCTCCGCTCAACAGAAAC	5100
Db	5176	GGGCTCTGATCTGTGCGAGCATGCCACAGCAAAAGTGTGTCATTTCTCCGCTCAACAGAAAC	5235
Qy	5101	CTCAGCAAACTCTGCATCCGGAAGAGATAGACCTTCAGAGCCCTCAGCTGTATCCAC	5160
Db	5236	CTCAGCAAACTCTGCATCCGGAAGAGATAGACCTTCAGAGCCCTCAGCTGTATCCAC	5295
Qy	5161	TTCAACAAATTACAGTATCCTCATTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAG	5220
Db	5296	TTCAACAAATTACAGTATCCTCATTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAG	5355

QY 5221 TACAGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGACACCTGCTGTGTTGCC 5280
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGACACCTGCTGTGTTGCC 5415
QY 5281 GCCTCTTCCACAGAGCTTCCCTCTCAATCGTGCAGGTGGAACAGCGCGAGGCGAGGAG 5340
Db 5416 GCCTCTTCCACAGAGCTTCCCTCTCAATCGTGCAGGTGGAACAGCGCGAGGCGAGGAG 5475
QY 5341 GAGTACTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGAGCTAGC 5400
Db 5476 GAGTACTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGAGCTAGC 5535
QY 5401 CGCACAGAGCATCTCAAGTGGAGTGCCTTACCTTTGGGCTTTGGCTTACAGAGAACCTTAT 5460
Db 5536 CGCACAGAGCATCTCAAGTGGAGTGCCTTACCTTTGGGCTTTGGCTTACAGAGAACCTTAT 5595
QY 5461 CTGTTTGTGACCACTTCAACTCACTCGAAGTAATGAGATCCAGGCGAGTCTCTCAGCA 5520
Db 5596 CTGTTTGTGACCACTTCAACTCACTCGAAGTAATGAGATCCAGGCGAGTCTCTCAGCA 5655
QY 5521 GGGACCCCTGCGCGAGCGTACCTGGACATCCGAAACCCGCGTACCTGGGCGCTGCCATT 5580
Db 5656 GGGACCCCTGCGCGAGCGTACCTGGACATCCGAAACCCGCGTACCTGGGCGCTGCCATT 5715
QY 5581 TCCTCAGGAGCGATTACTTGGCGTCCCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5640
Db 5716 TCCTCAGGAGCGATTACTTGGCGTCCCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5775
QY 5641 AAGGGAACCTCGTGAAGAGTCCGCGACTGAACACCCAGCGGCGCGTCCACCTCCCGC 5700
Db 5776 AAGGGAACCTCGTGAAGAGTCCGCGACTGAACACCCAGCGGCGCGTCCACCTCCCGC 5835
QY 5701 AGCAGCCCAACAGGAGGCGCAACCCAGTACAGTACAGAGACATCAACAGCGCGTGCC 5760
Db 5836 AGCAGCCCAACAGGAGGCGCAACCCAGTACAGTACAGAGACATCAACAGCGCGTGCC 5895
QY 5761 TCCAGCGCAGCGCCCGAGGCGCCAGCCAGCCAGCGAGAGCCAGCACACCCACCGC 5820
Db 5896 TCCAGCGCAGCGCCCGAGGCGCCAGCCAGCCAGCGAGAGCCAGCACACCCACCGC 5955
QY 5821 TACCGCGAGGCGGAGCCAGGCTGCGCAGGGAACAAGTCTCTGGCGCGCCCTTGGAGCGA 5880
Db 5956 TACCGCGAGGCGGAGCCAGGCTGCGCAGGGAACAAGTCTCTGGCGCGCCCTTGGAGCGA 6015
QY 5881 GAGAAGTCCCGCGCGAGTCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5940
Db 6016 GAGAAGTCCCGCGCGGATCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6075
QY 5941 GAAGACAGCAGGCGGCGCGTCTGCGGAGCGGTGAGGACCCCGCTGTCCAGGTG 6000
Db 6076 GAAGACAGCAGGCGGCGCGTCTGCGGAGCGGTGAGGACCCCGCTGTCCAGGTG 6135
QY 6001 AACAGGGAAGGCGAGAGTGC 6023
Db 6136 AACAGGCTCGGACCAAGTCTTC 6158

RESULT 8
AX166510
LOCUS AX166510 6159 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1 from Patent WO0138503.
ACCESSION AX166510
VERSION AX166510.1 GI:14546855
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plozman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S., Martinez, R.,
Planagan, P. and Clary, D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes

JOURNAL Patent: WO 0138503-A 1 31-MAY-2001;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source 1..6159
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 86.0%; Score 5650.4; DB 6; Length 6159;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 5943; Conservative 0; Mismatches 11; Indels 246; Gaps 4;
QY 19 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGGATCCTGGTCTGCTGTAACCCCAATT 78
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGGATCCTGGTCTGCTGTAACCCCAATT 60
QY 79 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 138
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
QY 139 CAGCAGATGTCTCTCTTTCCGAGAGGAGATATAGATCCCTCTTTTGTCTCTTTGAA 198
Db 121 CAGCAGATGTCTCTCTTTCCGAGAGGAGATATAGATCCCTCTTTTGTCTCTTTGAA 180
QY 199 GAATGCACTCAGCTGCTCTGATGAAGATTAAGCAGTGAAGCACTTTGTCGGA--G 255
Db 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAGCAGTGAAGCACTTTGTCGGAAGTG 240
QY 256 TATTCACACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGCAAGAGACTTTCGAAGTC 315
Db 241 TATTCACACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGCAAGAGACTTTCGAAGTC 300
QY 316 AGAAGTCTTGTAGTGTGTGCTACCTTTGCTGAGTGCAGTGGTGAAGAGAAAGCAACC 375
Db 301 AGAAGTCTTGTAGTGTGTGCTACCTTTGCTGAGTGCAGTGGTGAAGAGAAAGCAACC 360
QY 376 GGGGACATCTATGCTATGAAGTGAAGAGAGGCTTTATTTGGCCACGAGCAGGTT 435
Db 361 GGGGACATCTATGCTATGAAGTGAAGAGAGGCTTTATTTGGCCACGAGCAGGTT 420
QY 436 TCATTTTTCAGGAAGAGCGGAACATATATTCGAAGACAAAGCCCGTGGATCCCCCAA 495
Db 421 TCATTTTTCAGGAAGAGCGGAACATATATTCGAAGACAAAGCCCGTGGATCCCCCAA 480
QY 496 TTACAGTATGCCTTTCAGGACAAATATCAGCTTTATCTGATGGAGGAATATCAGCCTGGA 555
Db 481 TTACAGTATGCCTTTCAGGACAAATATCAGCTTTATCTGATGGAGGAATATCAGCCTGGA 540
QY 556 GGGGACTTGTGTCACCTTTTGAATAGATATGAGGACCCAGTTAGATGAAAACCTGATACAG 615
Db 541 GGGGACTTGTGTCACCTTTTGAATAGATATGAGGACCCAGTTAGATGAAAACCTGATACAG 600
QY 616 TTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTATCTGATGGGATACGTCAT 675
Db 601 TTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTATCTGATGGGATACGTCAT 660
QY 676 CGAGACATCAAGCCTGAGAACTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGAT 735
Db 561 CGAGACATCAAGCCTGAGAACTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGAT 720
QY 736 TTTGGATCTGCCGGAATAATGAATTCAAAACAAGTGTGAATGCCAAACTCCCGATTGGG 795
Db 721 TTTGGATCTGCCGGAATAATGAATTCAAAACAAGTGTGAATGCCAAACTCCCGATTGGG 780
QY 796 ACCCCAGATTACATGGCTCTCTGAGTGTGCTGATGAACGGGATGAAAAGGCACC 855
Db 781 ACCCCAGATTACATGGCTCTCTGAGTGTGCTGATGAACGGGATGAAAAGGCACC 840
QY 856 TACGGCCTGAGCTGTGATGCTGGTCACTAGTGGCGGTGATTCCTTATGAGATGATTTATGGG 915
Db 841 TACGGCCTGAGCTGTGATGCTGGTCACTGGTGTGCTGATTCCTTATGAGATGATTTATGGG 900

QY 916 AGATCCCCCTTCGAGAGGGAACCTCTGCGAGAAACCTTCAATTAACATTAATGAATTTCCAG 975
Db 901 AGATCCCCCTTCGAGAGGGAACCTCTGCGAGAAACCTTCAATTAACATTAATGAATTTCCAG 960
QY 976 CGGTTTTTGAATTTCCAGATGACCCCAAGTAGAGCAGTACATTTCTTGATCTGATTCAA 1035
Db 961 CGGTTTTTGAATTTCCAGATGACCCCAAGTAGAGCAGTACATTTCTTGATCTGATTCAA 1020
QY 1036 AGCTTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTGCTGCCATCTTTTC 1095
Db 1021 AGCTTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTGCTGCCATCTTTTC 1080
QY 1096 TTCTCTAAATTTGACTCGAACAATTCGTAATCTCTCTCCCTCCCTTCGTTCCACCCCTC 1155
Db 1081 TTCTCTAAATTTGACTCGAACAATTCGTAATCTCTCTCCCTCCCTTCGTTCCACCCCTC 1140
QY 1156 AAGTCTGACGATGACACCTCCAAATTTTGATGAACCCAGAGAAATTCGTTGGTTTCATCC 1215
Db 1141 AAGTCTGACGATGACACCTCCAAATTTTGATGAACCCAGAGAAATTCGTTGGTTTCATCC 1200
QY 1216 TCTCCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCGGTTTGTGGGTTT 1275
Db 1201 TCTCCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCGGTTTGTGGGTTT 1260
QY 1276 TCGTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTTGAGTCTGTGTGGGTCTGGAC 1335
Db 1261 TCGTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTTGAGTCTGTGTGGGTCTGGAC 1320
QY 1336 TCCCTGCCAAGACTAGCTCAATGGAAAGAACTTCTCATCAAAAAGCAAGAGCTACAA 1395
Db 1321 TCCCTGCCAAGACTAGCTCAATGGAAAGAACTTCTCATCAAAAAGCAAGAGCTACAA 1380
QY 1396 GACTCTCAGACAAAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCGGAGTG 1455
Db 1381 GACTCTCAGACAAAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCGGAGTG 1440
QY 1456 TCAGAGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGTGAAGCCCTCTGAGACTCAG 1515
Db 1441 TCAGAGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGTGAAGCCCTCTGAGACTCAG 1500
QY 1516 AGATCCCTCTGAGCAGGACCTGTGCTACCTATCACAGAAATGAGTAGCTTAAGCGA 1575
Db 1501 AGATCCCTCTGAGCAGGACCTGTGCTACCTATCACAGAAATGAGTAGCTTAAGCGA 1560
QY 1576 AGTTTGGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCGAGTTCTC 1635
Db 1561 AGTTTGGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCGAGTTCTC 1620
QY 1636 CATGATATCAGAGCAGGACCGGAAGCTTCAAGAAATCAAGAGCAGGAGTACAGGCT 1695
Db 1621 CATGATATCAGAGCAGGACCGGAAGCTTCAAGAAATCAAGAGCAGGAGTACAGGCT 1680
QY 1696 CAAGTGAAGAAATGAGTTGATGAATCACTTGGAAAGAGGATCTTGTCTCAGCAAGA 1755
Db 1681 CAAGTGAAGAAATGAGTTGATGAATCACTTGGAAAGAGGATCTTGTCTCAGCAAGA 1740
QY 1756 AGACGAGTGAATCTTACGAATCTGAGCTGAGAGCTCTCGGTTGCTGCTGAAGATTC 1815
Db 1741 AGACGAGTGAATCTTACGAATCTGAGCTGAGAGCTCTCGGTTGCTGCTGAAGATTC 1800
QY 1816 AAGCGAAACGACAGAAATGTGAGCAATAACTTGTGAAGCTTAAGGATCAAGGAGGCT 1875
Db 1801 AAGCGAAACGACAGAAATGTGAGCAATAACTTGTGAAGCTTAAGGATCAAGGAGGCT 1860
QY 1876 GAAGTGGGAGAAATATCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAG 1935
Db 1861 GAAGTGGGAGAAATATCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAG 1920
QY 1936 GAGCTCCAGAGAACTGGAGAGGCT----- 1962
Db 1921 GAGCTCCAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGTG 1980
QY 1963 -----GCAAAGGAGCGAGCGGAGGAGCTGGAGAGGCTGAGAAACCGA 2007

Db 1981 CAGAAATATCCCGCAAGGCAAGAGCGAGCCGAGAGGAGCTGAGAAAGCTTCAGAACCGA 2040
QY 2008 GAGGATTTCTTCTGAGGCGATCAGAAAGAAAGCTGTGGAAGCTGAGGAACGCGCATTTCT 2067
Db 2041 GAGGATTTCTTCTGAGGCGATCAGAAAGAAAGCTGTGGAAGCTGAGGAACGCGCATTTCT 2100
QY 2068 CTGGAGAAACAGGTAAAGAGACTAGAGACCATGAGCGGTAGAGAAAACAGACTGAAGAT 2127
Db 2101 CTGGAGAAACAGGTAAAGAGACTAGAGACCATGAGCGGTAGAGAAAACAGACTGAAGAT 2160
QY 2128 GACTTCCAGACAAATCCCAACAGATCCAGCAGATGCTGATAAATTTCTGGAGCTCGAA 2187
Db 2161 GACTTCCAGACAAATCCCAACAGATCCAGCAGATGCTGATAAATTTCTGGAGCTCGAA 2220
QY 2188 GAGAAACATCTGGGAGGCCCCAAGTCTCAGCCAGCACCTAGAAAGTGCACCTGAAACAGAAA 2247
Db 2221 GAGAAACATCTGGGAGGCCCCAAGTCTCAGCCAGCACCTAGAAAGTGCACCTGAAACAGAAA 2280
QY 2248 GAGCAGCACTATGAGGAAAGATTAAGTGTTCGCAATCAGATTAAGAAAGACCTGGCT 2307
Db 2281 GAGCAGCACTATGAGGAAAGATTAAGTGTTCGCAATCAGATTAAGAAAGACCTGGCT 2340
QY 2308 GACAAAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAAGGC 2367
Db 2341 GACAAAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAAGGC 2400
QY 2368 AAAATTTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTTCCAAGATCAGATCCCTG 2427
Db 2401 AAAATTTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTTCCAAGATCAGATCCCTG 2460
QY 2428 GAA CAGAGGATTTGGAACTGTCTGAAAGCCAAATAAACTTCAGCAAAATAGCAGTCTTTT 2487
Db 2461 GAA CAGAGGATTTGGAACTGTCTGAAAGCCAAATAAACTTCAGCAAAATAGCAGTCTTTT 2520
QY 2488 ACCCAAAGGAACTGAAGGCCCAAGAGAGATGATTTTCTGAACTCAGGCAACAGAAATTT 2547
Db 2521 ACCCAAAGGAACTGAAGGCCCAAGAGAGATGATTTTCTGAACTCAGGCAACAGAAATTT 2580
QY 2548 TACCTGAGACACAGGCTGGGAGTTGAGGCCCGAGAACCGAAACCTGAGGAGCAGCTG 2607
Db 2581 TACCTGAGACACAGGCTGGGAGTTGAGGCCCGAGAACCGAAACCTGAGGAGCAGCTG 2640
QY 2608 GAGAAAGTCAAGCCACCAAGACCAAGTGTGCAAGAAATCGGCTGTGGAACCTGGAGACAAGA 2667
Db 2641 GAGAAAGTCAAGCCACCAAGACCAAGTGTGCAAGAAATCGGCTGTGGAACCTGGAGACAAGA 2700
QY 2668 TTGCGGAGGTCAGTCTAGAGCAGAGGAGAGAACTGGAGCTCAAGCGCCAGCTCACA 2727
Db 2701 TTGCGGAGGTCAGTCTAGAGCAGAGGAGAGAACTGGAGCTCAAGCGCCAGCTCACA 2760
QY 2728 GAGCTACAGCTCTCCCTGCGAGGCGGAGTCAAGTTGACAGCCCTGCGAGGCTGCACGG 2787
Db 2761 GAGCTACAGCTCTCCCTGCGAGGCGGAGTCAAGTTGACAGCCCTGCGAGGCTGCACGG 2820
QY 2788 GCGGCCCTGAGAGCAGCTTCCGAGGCCAAGACAGAGCTGGAAGACACAGCAGAA 2847
Db 2821 GCGGCCCTGAGAGCAGCTTCCGAGGCCAAGACAGAGCTGGAAGACACAGCAGAA 2880
QY 2848 GCTCAAGAGGAGATTCAGGCACTCACGGCACATAGAGATGAAATCCAGCCCAATTTGAT 2907
Db 2881 GCTCAAGAGGAGATTCAGGCACTCACGGCACATAGAGATGAAATCCAGCCCAATTTGAT 2940
QY 2908 GCTCTTCGTAACAGCTGTA CTGTAATCAGAGACTGGAGGAGCAGCTAAACAGAGTACC 2967
Db 2941 GCTCTTCGTAACAGCTGTA CTGTAATCAGAGACTGGAGGAGCAGCTAAACAGAGTACC 3000
QY 2968 GAGGACAACTGAACCTCAACACCAAACTTCTACTTTGCCAAAACAACTCGATGAGGCT 3027
Db 3001 GAGGACAACTGAACCTCAACACCAAACTTCTACTTTGCCAAAACAACTCGATGAGGCT 3060
QY 3028 TCTGGCCCAACGACGAGATTGTACAACTCGAAGTGAAGTGAACCATTTCTCCCGCGGAG 3087

Db 3061 TCTGGCGCCAAACGACGAGATTGTCAAACTGCGAAGTGAAGTGGACCATCTCCGCGGGAG 3120
QY 3088 ATCAGGGAACGAGATGACGAGCTTACAGCCAGAAAGCAAAACCATGAGGCTCTGAAGACC 3147
Db 3121 ATCAGGAACGAGATGACGAGCTTACAGCCAGAAAGCAAAACCATGAGGCTCTGAAGACC 3180
QY 3148 AGTGACCATGCTGGAGGAAACAGGTCATGGAATTTGGAGGCCCTAAACGATGAGTGCTA 3207
Db 3181 AGTGACCATGCTGGAGGAAACAGGTCATGGAATTTGGAGGCCCTAAACGATGAGTGCTA 3240
QY 3208 GAAAAAGAGCGCAGTGGAGGCCCTGGAGGCGCTCTGGGTGATGAGAAATCCAGATT 3267
Db 3241 GAAAAAGAGCGCAGTGGAGGCCCTGGAGGCGCTCTGGGTGATGAGAAATCCAGATT 3300
QY 3268 GAGTGTCCGGTTCGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGA 3327
Db 3301 GAGTGTCCGGTTCGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGA 3360
QY 3328 GCCGATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGCACAAAG 3387
Db 3361 GCCGATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGCACAAAG 3420
QY 3388 GCTGAGATTCTCGCTCTGCAGCAGGCTCTCAAGAGCAGAGCTCAAGGCCGAGAGCCCTC 3447
Db 3421 GCTGAGATTCTCGCTCTGCAGCAGGCTCTCAAGAGCAGAGCTCAAGGCCGAGAGCCCTC 3480
QY 3448 TCTGACAGCTCAATGACCTGGAGAGAGGATGCTATGCTTGAATGAATGAATGCCCGAAGC 3507
Db 3481 TCTGACAGCTCAATGACCTGGAGAGAGGATGCTATGCTTGAATGAATGAATGCCCGAAGC 3540
QY 3508 TTACAGCAGAAGCTGGAGACTCAACGAGAGCTCAACAGAGAGGCTCTGGAAGAGCAGGCC 3567
Db 3541 TTACAGCAGAAGCTGGAGACTCAACGAGAGCTCAACAGAGAGGCTCTGGAAGAGCAGGCC 3600
QY 3568 AAATTACAGCAGATGGAACCTGAGAAATCAATTTCCGTCTGACTCAAGGACTG 3627
Db 3601 AAATTACAGCAGATGGAACCTGAGAAATCAATTTCCGTCTGACTCAAGGACTG 3660
QY 3628 CAAGAAGCTCTAGATCGGGCTGATCTACTGAGCAGAGAGAGTGAAGTGGAGTATCAG 3687
Db 3661 CAAGAAGCTCTAGATCGGGCTGATCTACTGAGCAGAGAGAGTGAAGTGGAGTATCAG 3720
QY 3688 CTGGAACCAATTCAGGTTCTCTATTCTCATGAAAGGTTGAAATGGAAGGCACTATTCT 3747
Db 3721 CTGGAACCAATTCAGGTTCTCTATTCTCATGAAAGGTTGAAATGGAAGGCACTATTCT 3780
QY 3748 CAACAAACCAACTCATTTGATTTCTGAGCCAAATGGACCAACCTGCTAAAGAA 3807
Db 3781 CAACAAACCAACTCATTTGATTTCTGAGCCAAATGGACCAACCTGCTAAAGAA 3840
QY 3808 AAGGGTTTATTAGTCGAGGAAAGGAGACCTGCTTTACCCACACAGGTTCCCTCTGCAG 3867
Db 3841 A-----AGTTCTCTGCAG 3855
QY 3868 TACAATGAGCTGAAGCTGGCCCTGAGAGAGGAAAGCTGCTGTGACAGCTAGAGAA 3927
Db 3856 TACAATGAGCTGAAGCTGGCCCTGAGAGAGGAAAGCTGCTGTGACAGCTAGAGAA 3915
QY 3928 GCCCTTCAGAGACCCGAGCTCAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAA 3987
Db 3916 GCCCTTCAGAGACCCGAGCTCAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAA 3975
QY 3988 GCAACGGACCCACCACTCCACGCGCAGCCACCGCGAGGAGCAGATGCCATGTCC 4047
Db 3976 GCAACGGACCCACCACTCCACGCGCAGCCACCGCGAGGAGCAGATGCCATGTCC 4035
QY 4048 GGCATGTGCGGTGCCGAGAGCAGCCAGTGCATGAGGCTGCTGCGCCCGGCATCC 4107
Db 4036 GGCATGTGCGGTGCCGAGAGCAGCCAGTGCATGAGGCTGCTGCGCCCGGCATCC 4095
QY 4108 AGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGATG 4167
Db 4096 AGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGATG 4155

QY 4168 CACCACAATATTCTCTCACCAGTTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGCT 4227
Db 4156 CACCACAATATTCTCTCACCAGTTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGCT 4215
QY 4228 GTGTGTCTGATACCGTGCACTTTGGAGCCAGGCATCCAAATGTCTCGAATGTCAAGTG 4287
Db 4216 GTGTGTCTGATACCGTGCACTTTGGAGCCAGGCATCCAAATGTCTCGAATGTCAAGTG 4275
QY 4288 ATGTGTCAACCCCAAGTCTCCACGTGCTTGCAGCCACCTGCGGCTTGCCTGCTGAATAT 4347
Db 4276 ATGTGTCAACCCCAAGTCTCCACGTGCTTGCAGCCACCTGCGGCTTGCCTGCTGAATAT 4335
QY 4348 GCCACACATTTCAACCGAGGCTTTCTGCCGTGACAAATGAATGAACTCCCAAGTCTCCAGACC 4407
Db 4336 GCCACACATTTCAACCGAGGCTTTCTGCCGTGACAAATGAATGAACTCCCAAGTCTCCAGACC 4395
QY 4408 AAGGAGCCACGACGAGCTTGCACCTTGGAGGCTGGATGAAGTGGCCAGCAATACAAA 4467
Db 4396 AAGGAGCCACGACGAGCTTGCACCTTGGAGGCTGGATGAAGTGGCCAGCAATACAAA 4455
QY 4468 CGAGGACAGCAAGGCTGGGACAGGAATGATCTCTCGAGGGATCAAAAGTCTCTCAT 4527
Db 4456 CGAGGACAGCAAGGCTGGGACAGGAATGATCTCTCGAGGGATCAAAAGTCTCTCAT 4515
QY 4528 TATGACAAATGAAGCCAGAGAGCTGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCCTT 4587
Db 4516 TATGACAAATGAAGCCAGAGAGCTGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCCTT 4575
QY 4588 CCCACCGGATGTATCTATTCTATGTTGCTGCTTCCGAACTCGCAATACAGCC 4647
Db 4576 CCCACCGGATGTATCTATTCTATGTTGCTGCTTCCGAACTCGCAATACAGCC 4635
QY 4648 AAAGCAGA----- 4655
Db 4636 AAAGCAGATGTCCTCCATACATACATACTCAAGATGGAATCTCACCCGACACCACTGCTGCC 4695
QY 4656 ----- 4655
Db 4696 GGGAGAACCTCTACTTGTGCTCCAGCTTCCCTGACAAACAGCGCTGGGTACCGCC 4755
QY 4656 -----AAAAGCAGAAGCTGATGCTAAA 4677
Db 4756 TTAGATCAGTTGTGCGAGGTGGAGAGTTTCTAGGGAAGAAAGCAGAGCTGATCTAAA 4815
QY 4678 CTGCTTGGAACTCCCTGCTGAAACTGGAGGTGATGACCGTCTAGACATGAATGACG 4737
Db 4816 CTGCTTGGAACTCCCTGCTGAAACTGGAGGTGATGACCGTCTAGACATGAATGACG 4875
QY 4738 CTGCTTCAAGTACACAGGTGTTGTTGGTGCCACCGAGGAGGCTCTACGCGCTGAAT 4797
Db 4876 CTGCTTCAAGTACACAGGTGTTGTTGGTGCCACCGAGGAGGCTCTACGCGCTGAAT 4935
QY 4798 GTCTTGAAGAACTCCCTTAACCCATGTCCAGGAATTTGGAGAGCTCTTCCAAATTTATAT 4857
Db 4936 GTCTTGAAGAACTCCCTTAACCCATGTCCAGGAATTTGGAGAGCTCTTCCAAATTTATAT 4995
QY 4858 ATCAAGGCTGAGAGAGCTACTCATGATAGAGGAGAGAGCGGACCTGCTCTGTG 4917
Db 4996 ATCAAGGCTGAGAGAGCTACTCATGATAGAGGAGAGAGCGGACCTGCTCTGTG 5055
QY 4918 GAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATC 4977
Db 5056 GAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATC 5115
QY 4978 TCACCCAACTTTTGAAGTGTCAAGGCTGCCACTGTTTGGGAGGAGAGATGAG 5037
Db 5116 TCACCCAACTTTTGAAGTGTCAAGGCTGCCACTGTTTGGGAGGAGAGATGAG 5175
QY 5038 AAGCGGCTCTGATCTGTGAGCAGCTGCCAGCAAGTGTCTCATTTCTCGCTACACGAA 5097
Db 5176 AAGCGGCTCTGATCTGTGAGCAGCTGCCAGCAAGTGTCTCATTTCTCGCTACACGAA 5235

QY 5098 AACCTCAGCAATACATGCTCGGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATC 5157
Db 5236 AACCTCAGCAATACATGCTCGGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATC 5295
QY 5158 CACTTCACCAATACAGTATCTCTCATTTGGAACCAATAAATTTCTAGCAATGACATGAAG 5217
Db 5296 CACTTCACCAATACAGTATCTCTCATTTGGAACCAATAAATTTCTAGCAATGACATGAAG 5355
QY 5218 CAGTACACGCTCAGCAATCTCTGGAATAGATGACATCTCTGCGACCTCTGCTGTTT 5277
Db 5356 CAGTACACGCTCAGCAATCTCTGGAATAGATGACATCTCTGCGACCTCTGCTGTTT 5415
QY 5278 GCCGCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGAGCGA 5337
Db 5416 GCCGCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGAGCGA 5475
QY 5338 GAGAGTACTTGTGTTTTCACGAATTTGGAGTCTTCTGGAATTTTACGGAAGAGT 5397
Db 5476 GAGAGTACTTGTGTTTTCACGAATTTGGAGTCTTCTGGAATTTTACGGAAGAGT 5535
QY 5398 AGCGGACAGACGATCTCAAGTGGAGTCTGCTTACCTTTGCGCTTACGAGAACCC 5457
Db 5536 AGCGGACAGACGATCTCAAGTGGAGTCTGCTTACCTTTGCGCTTACGAGAACCC 5595
QY 5458 TATCTGTTGTGACCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCTCA 5517
Db 5596 TATCTGTTGTGACCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCTCA 5655
QY 5518 GCAGGACCCCTCCGAGGTACCTGGACATCCCGAACCCCGCTACCTGGGCCCTGCC 5577
Db 5656 GCAGGACCCCTCCGAGGTACCTGGACATCCCGAACCCCGCTACCTGGGCCCTGCC 5715
QY 5578 ATTTCTCTCAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGGTCAATTGC 5637
Db 5716 ATTTCTCTCAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGGTCAATTGC 5775
QY 5638 TGAAGGAAACCTCTGTAAGAGTCTGGGACCTGAACACCCCGGGCCCGTCCACCTCC 5697
Db 5776 TGAAGGAAACCTCTGTAAGAGTCTGGGACCTGAACACCCCGGGCCCGTCCACCTCC 5835
QY 5698 CGCAGACGCCCCAACAGCGAGGCCACCCACCGTACAGAGCACATCAACGAGCGCGTG 5757
Db 5836 CGCAGACGCCCCAACAGCGAGGCCACCCACCGTACAGAGCACATCAACGAGCGCGTG 5895
QY 5758 GCCTCCAGCCAGCGCGCCCGGAGGCCAGCCACCGGAGAGAGGAGAGAGAGAGAGAGAG 5817
Db 5896 GCCTCCAGCCAGCGCGCCCGGAGGCCAGCCACCGGAGAGAGGAGAGAGAGAGAGAGAG 5955
QY 5818 CGCTACCGAGGCGGACCGAGCTGCGCAGGAGCAAGTCTCTGCGCCCGCCCTGGAG 5877
Db 5956 CGTACCGAGGCGGACCGAGCTGCGCAGGAGCAAGTCTCTGCGCCCGCCCTGGAG 6015
QY 5878 CGAGAGAGTCCCGCCCGGAGTCTCAGCAACCGGAGAGAGCGGTCCCGCGGAGGCTG 5937
Db 6016 CGAGAGAGTCCCGCCCGGAGTCTCAGCAACCGGAGAGAGCGGTCCCGCGGAGGCTG 6075
QY 5938 TTTGAGACAGCAGCGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGCTCCAG 5997
Db 6076 TTTGAGACAGCAGCGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGCTCCAG 6135
QY 5998 GTGAACAGGGAAGAGGGCA 6017
Db 6136 GTGAACAGGGAAGAGGGCA 6155

RESULT 9
AX503780
LOCUS AX503780 6189 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 10 from Patent WO0226826.
ACCESSION AX503780
VERSION AX503780.1 GI:23385965
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Gerlach, V.L., Macdougall, J.R., Smithson, G., Millet, I., Stone, D., Gunther, E., Ellerman, K., Grosse, W.M., Alsobrook, J.P., Lepley, D.M., Burgess, C.E., Padigaru, M., Kekuda, R., Spytek, K.A., Leach, M.D. and Shimkets, R.A.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0226826-A 10 04-APR-2002;
FEATURES Curagen Corporation (US)
source location/Qualifiers
1. 6189
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 85.7%; Score 5631; DB 6; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;
QY 19 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 78
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 60
QY 79 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCCTTTATGACTCAA 138
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCCTTTATGACTCAA 120
QY 139 CAGCAGATGCTCTCTCTTCCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198
Db 121 CAGCAGATGCTCTCTCTTCCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 199 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCACTTTGTCGGAAGTAT 258
Db 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCACTTTGTCGGAAGTAT 240
QY 259 TCCGACACCATAGCTAGTTTACAGGAGCTCCAGCCTTCGGCAAGGAGCTTCGAAGTCAGA 318
Db 241 TCCGACACCATAGCTAGTTTACAGGAGCTCCAGCCTTCGGCAAGGAGCTTCGAAGTCAGA 300
QY 319 AGTCTGTAGTGTGCTCACTTTGCTGAAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG 378
Db 301 AGTCTGTAGTGTGCTCACTTTGCTGAAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG 360
QY 379 GACATCTATGCTAAGTGTAGTGAAGAAAGAGCTTTATGCGCCAGGAGAGGTTTCA 438
Db 361 GACATCTATGCTAAGTGTAGTGAAGAAAGAGCTTTATGCGCCAGGAGAGGTTTCA 420
QY 439 TTTTGTGAGAGAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 498
Db 421 TTTTGTGAGAGAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480
QY 499 CAGTATGCTTTTACGACAAATAATCACCTTTATCTGATGAGGAAATATCAGCTGGAGGG 558
Db 481 CAGTATGCTTTTACGACAAATAATCACCTTTATCTGATGAGGAAATATCAGCTGGAGGG 540
QY 559 GACTTGTGTCACCTTTTGAATAGATATGAGACCCAGTTAGATGAAACCTGATACAGTTT 618
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGACCCAGTTAGATGAAACCTGATACAGTTT 600
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCTCTGATGGGATACGTCGATCGA 678
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCTCTGATGGGATACGTCGATCGG 660
QY 679 GACATCAAGCTGAGACATTTCTGTTGACCGCAGGACACATCAAGCTGGTGGATTTT 738
Db 661 GACATCAAGCTGAGACATTTCTGTTGACCGCAGGACACATCAAGCTGGTGGATTTT 720
QY 739 GGATCTGCGCGGAAATGAATTCAAAACAGATGGTGAATGCCAAACTCCCGGATGGGACC 798
Db 721 GGATCTGCGCGGAAATGAATTCAAAACAA---GGTGAATGCCAAACTCCCGGATGGGACC 777

QY 799 CCAGATTACATGGCTCCCTGAAAGTGTGATGTAACCGGGATGGAAGGCACTTAC 858
DB 778 CCAGATTACATGGCTCCCTGAAAGTGTGATGTAACCGGGATGGAAGGCACTTAC 837
QY 859 GGCCTGGAGTGTGACTGGTGGTCACTGGGCGGTGATGCTATGAGATGATTTATGGGAGA 918
DB 838 GGCCTGGAGTGTGACTGGTGGTCACTGGGCGGTGATGCTATGAGATGATTTATGGGAGA 897
QY 919 TCCCTCTGCGAGAGGGAACCTCTGCCAGAACCTTCAATAACATPATGAATTCAGCGG 978
DB 898 TCCCTCTGCGAGAGGGAACCTCTGCCAGAACCTTCAATAACATPATGAATTCAGCGG 957
QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTGGACGTGACTTCTTGATCTGATTCARAGC 1038
DB 958 TTTTGTGAATTTCCAGATGACCCCAAGTGGACGTGACTTCTTGATCTGATTCARAGC 1017
QY 1039 TTGTTGTGGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTGCCATCCTTTCTC 1098
DB 1018 TTGTTGTGGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTGCCATCCTTTCTC 1077
QY 1099 TCTAAATTTGACTGGAACAACTTCGTAACTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1158
DB 1078 TCTAAATTTGACTGGAACAACTTCGTAACTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1137
QY 1159 TCTGACGATGACACCTCCAAATTTTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1218
DB 1138 TCTGACGATGACACCTCCAAATTTTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1197
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGCTTTGTGGGTTTTCG 1278
DB 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGCTTTGTGGGTTTTCG 1257
QY 1279 TACAGCAAGGCACTGGGATCTTCTGTAGTCTGTAGTCTGTGTGTCGGGTCGACTCC 1338
DB 1258 TACAGCAAGGCACTGGGATCTTCTGTAGTCTGTAGTCTGTGTGTCGGGTCGACTCC 1317
QY 1339 CTGCGCAAGACTAGCTCCATGGAAGAACTTCTCATCAAGCAAGAAAGACTCAAGAC 1398
DB 1318 CTGCGCAAGACTAGCTCCATGGAAGAACTTCTCATCAAGCAAGAAAGACTCAAGAC 1377
QY 1399 TCTAGGCAAGTGTCAAGATGAGCAGGAAATGACCCGTTACATCGAGAGTGTCA 1458
DB 1378 TCTAGGCAAGTGTCAAGATGAGCAGGAAATGACCCGTTACATCGAGAGTGTCA 1437
QY 1459 GAGTGGAGGCTGTCTTAGTCAAGAGGAGTGGAGCTGAAAGGCTCTGAGACTCAGAGA 1518
DB 1438 GAGTGGAGGCTGTCTTAGTCAAGAGGAGTGGAGCTGAAAGGCTCTGAGACTCAGAGA 1497
QY 1519 TCCCTCTGGAGCAGGACCTTGCTACCTACATCAGAAATGCGATGAGTTAAAGGAAGT 1578
DB 1498 TCCCTCTGGAGCAGGACCTTGCTACCTACATCAGAAATGCGATGAGTTAAAGGAAGT 1557
QY 1579 TTGGAGCAAGCAGGATGGAGTGTCCAGAGGAGTGAACAGCACTGACGCTTCTCCAT 1638
DB 1558 TTGGAGCAAGCAGGATGGAGTGTCCAGAGGAGTGAACAGCACTGACGCTTCTCCAT 1617
QY 1639 GATATCAGAGCAGAGCCGGAAGTCTCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1698
DB 1618 GATATCAGAGCAGAGCCGGAAGTCTCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677
QY 1699 GTGGAAGAAATCAGGTTGATGATGAATCAGTTGGAGAGGATCTTGCTCAGCAGAGA 1758
DB 1678 GTGGAAGAAATCAGGTTGATGATGAATCAGTTGGAGAGGATCTTGCTCAGCAGAGA 1737
QY 1759 CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTGCTGAAGAAATTCAG 1818
DB 1738 CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTGCTGAAGAAATTCAG 1797
QY 1819 CGGAAAGCAGAGATGTACGATTAACCTGTTGAAGGCTTAAGGATCAAGGAGCGCTGAA 1878
DB 1798 CGGAAAGCAGAGATGTACGATTAACCTGTTGAAGGCTTAAGGATCAAGGAGCGCTGAA 1857

QY 1879 GTGGGAGAAATATCGGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938
DB 1858 GTGGGAGAAATATCGGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1917
QY 1939 CTCACAGAGAAACTGGAGAGGCTG----- 1963
DB 1918 CTCACAGAGAAACTGGAGAGGCTGTAAGAGCCAGCAGCGAGGCCAGCGAGCTGCTGAG 1977
QY 1964 -----CAAGAGCGAGCGAGAGGAGCTGGAGAGCTGGAGAGCTGCAGAACCGAGAG 2010
DB 1978 AATATCCGCGAGGCAAGAGGAGCGAGCGAGGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2037
QY 2011 GATTCTTCTGAAGGCATCAGAAAGAGTCTGTTGAAGTCTGAGAAACCGCCCATTTCTCTG 2070
DB 2038 GATTCTTCTGAAGGCATCAGAAAGAGTCTGTTGAAGTCTGAGAAACCGCCCATTTCTCTG 2097
QY 2071 GAGAACCAAGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2130
DB 2098 GAGAACCAAGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2157
QY 2131 ATCCAGACAAATCCCAACAGATCCAGCAGCTAGAGTGTGATTAATTTCTGGAGCTCGAGAG 2190
DB 2158 ATCCAGACAAATCCCAACAGATCCAGCAGCTAGAGTGTGATTAATTTCTGGAGCTCGAGAG 2217
QY 2191 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGTGATTAATTTCTGGAGCTCGAGAG 2250
DB 2218 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGTGATTAATTTCTGGAGCTCGAGAG 2277
QY 2251 CAGCACTATGAGAAAGATTAAGTGTGGAACAATCAGATAAGAAAGACCTGGCTGAC 2310
DB 2278 CAGCACTATGAGAAAGATTAAGTGTGGAACAATCAGATAAGAAAGACCTGGCTGAC 2337
QY 2311 AAGGAGACACTGAGAAACATGATGAGAGACAGAGAGGAGGCCCATGAGAAAGGCGAAA 2370
DB 2338 AAGGAGACACTGAGAAACATGATGAGAGACAGAGAGGAGGCCCATGAGAAAGGCGAAA 2397
QY 2371 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGATTTCCAGATTCAGATCCCTGGAA 2430
DB 2398 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGATTTCCAGATTCAGATCCCTGGAA 2457
QY 2431 CAGAGGATTTGTGAACCTGCTGAAGCCATAAATTTGACGAAATAGCAGTCTTTTACC 2490
DB 2458 CAGAGGATTTGTGAACCTGCTGAAGCCATAAATTTGACGAAATAGCAGTCTTTTACC 2517
QY 2491 CAAAGGAACATGAGGCCCAAGAGAGATGATTTCTGAATTCAGGCAACAGAAATTTTAC 2550
DB 2518 CAAAGGAACATGAGGCCCAAGAGAGATGATTTCTGAATTCAGGCAACAGAAATTTTAC 2577
QY 2551 CTGGAGACAGAGCTGGGAAGTTGGAGGCCAGAACGGAATCTGGAGGAGCAGCTGGAG 2610
DB 2578 CTGGAGACAGAGCTGGGAAGTTGGAGGCCAGAACGGAATCTGGAGGAGCAGCTGGAG 2637
QY 2611 AAGATCAGCCCAAGACACAGTGAACAAGATTCGGCTGCTGGAACTGGAGACAAAGTTG 2670
DB 2638 AAGATCAGCCCAAGACACAGTGAACAAGATTCGGCTGCTGGAACTGGAGACAAAGTTG 2697
QY 2671 CGGAGGTGATCTTAGAGACAGAGAGCAGAACTCGAGCTCAAGCGCAGCTCAGAG 2730
DB 2698 CGGAGGTGATCTTAGAGACAGAGAGCAGAACTCGAGCTCAAGCGCAGCTCAGAG 2757
QY 2731 CTACAGCTCTCCCTGAGAGGCGGAGTCACTGTCACAGCCCTGACAGGCTGCACGCGG 2790
DB 2758 CTACAGCTCTCCCTGAGAGGCGGAGTCACTGTCACAGCCCTGACAGGCTGCACGCGG 2817
QY 2791 GCCCTGGAGAGCCAGCTTCGCCAGGGCAAGACAGAGCTTGAAGAGACCAACAGCAGAGCT 2850
DB 2818 GCCCTGGAGAGCCAGCTTCGCCAGGGCAAGACAGAGCTTGAAGAGACCAACAGCAGAGCT 2877
QY 2851 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATTCAGCGCAATTTGATGCT 2910
DB 2878 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATTCAGCGCAATTTGATGCT 2937
QY 2911 CTTTCGTAACAGCTGTACTGTAAATCAGACAGCCTGGAGGAGCAGCTAAACCAAGCTGACCGAG 2970

Qy	4981	CCCAACAATTTTGAAGCTGTCAAAGGCTGCCACTTGTTTGGGCAGGCAAGAATTGAGAAC	5041
Db	5113	CCCAACAATTTTGAAGCTGTCAAAGGCTGCCACTTGTTTGGGCAGGCAAGAATTGAGAAC	5172
Qy	5041	GGGCTCTGCATCTGTGCAGCCATGCCAGAAAGTCTGCTAATCTCCGCTACACGAAAC	5100
Db	5173	GGGCTCTGCATCTGTGCAGCCATGCCAGAAAGTCTGCTAATCTCCGCTACACGAAAC	5232
Qy	5101	CTCAGCAAAATCTGCATCCGGAAAGAGATAGAGACCTCAGAGCCCTGCTATCCAC	5160
Db	5233	CTCAGCAAAATCTGCATCCGGAAAGAGATAGAGACCTCAGAGCCCTGCTATCCAC	5292
Qy	5161	TTACACCAATACAGTATCCTCATTTGGAAACCAATAAATTTACGAAATCGACATGAAGCAG	5220
Db	5293	TTACACCAATACAGTATCCTCATTTGGAAACCAATAAATTTACGAAATCGACATGAAGCAG	5352
Qy	5221	TACACGCTCAGGAAATTCCTGGATAAGAAATGACCAATTCCTTGGCACCTGCTGTGTTGCC	5280
Db	5353	TACACGCTCAGGAAATTCCTGGATAAGAAATGACCAATTCCTTGGCACCTGCTGTGTTGCC	5412
Qy	5281	GCCTCTTCCAAACAGCTTCCTCTGCTCAATCGTGCAGGTGAACAGCGGACGGCAGCAGAG	5340
Db	5413	GCCTCTTCCAAACAGCTTCCTCTGCTCAATCGTGCAGGTGAACAGCGGACGGCAGCAGAG	5472
Qy	5341	GAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTTCGTGGAATTTCTACGGAACAGCTAGC	5400
Db	5473	GAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTTCGTGGAATTTCTACGGAACAGCTAGC	5532
Qy	5401	CGCACAGACGATCTCAAGTGGAGTCGTTTACCTTTTGGCCTTTCCTTACAGAGAACCTAT	5460
Db	5533	CGCACAGACGATCTCAAGTGGAGTCGTTTACCTTTTGGCCTTTCCTTACAGAGAACCTAT	5592
Qy	5461	CTGTTTGTGACCCACTTCAACTCAGTGAAGTAAATGAGATCCAGCAGCAGCTCCTCAGCA	5520
Db	5593	CTGTTTGTGACCCACTTCAACTCAGTGAAGTAAATGAGATCCAGCAGCAGCTCCTCAGCA	5652
Qy	5521	GGGACCCCTGCCGAGGCTACCTGGACATCCCGAAACCCGCGCTACTTGGGCCCTGCATT	5580
Db	5653	GGGACCCCTGCCGAGGCTACCTGGACATCCCGAAACCCGCGCTACTTGGGCCCTGCATT	5712
Qy	5581	TCCTCAGGACGATTTACTTTGGGTCTCTATACAGGATAAATAGGGTCAATTTGCTGC	5640
Db	5713	TCCTCAGGACGATTTACTTTGGGTCTCTATACAGGATAAATAGGGTCAATTTGCTGC	5772
Qy	5641	AAGGAAACCTCGTGAAGAGTCGGGACATGAAACCAACCGGGGCGCTCCACTCCCGC	5700
Db	5773	AAGGAAACCTCGTGAAGAGTCGGGACATGAAACCAACCGGGGCGCTCCACTCCCGC	5832
Qy	5701	AGCAGCCCCAAAGCGAGGCCACCAACCTAACAAGAGCACATCAACAGCGCTGGCC	5760
Db	5833	AGCAGCCCCAAAGCGAGGCCACCAACCTAACAAGAGCACATCAACAGCGCTGGCC	5892
Qy	5761	TCAGCCACAGCGCGCCGAAGCGCCAGCCACCGCGAGAGCAGCACACCCACCGC	5820
Db	5893	TCAGCCACAGCGCGCCGAAGCGCCAGCCACCGCGAGAGCAGCACACCCACCGC	5952
Qy	5821	TACCGAGGGGGGACCGAGCTCGCAGGGACAAAGTCTCCTGGCCGCCCTTGGAGCGA	5880
Db	5953	TACCGAGGGGGGACCGAGCTCGCAGGGACAAAGTCTCCTGGCCGCCCTTGGAGCGA	6012
Qy	5881	GAGAGTCCCCCGCGGATGCTCAGCACCGGAGAGAGCGGTCCCCCGGGAGGCTGTTT	5940
Db	6013	GAGAGTCCCCCGCGGATGCTCAGCACCGGAGAGAGCGGTCCCCCGGGAGGCTGTTT	6072
Qy	5941	GAAGACAGCAGCAGGGGCGGCTTGCCTGCGGAGCCGTGAGGACCCCGCTGTGCCAGGTG	6000
Db	6073	GAAGACAGCAGCAGGGGCGGCTTGCCTGCGGAGCCGTGAGGACCCCGCTGTGCCAGGTG	6132
Qy	6001	AACAGGGAAGAGGCGAGATGCTCTCAAGTTTTCACGGTTTAACTGTCACT	6055
Db	6133	AACAGGTTGGGACAGCTCTTTCAGTATAAATCTCAGCCAGAAAAACCACTCCT	6187

RESULT 10	AX503778	6201 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX503778	6201 bp	DNA	linear	PAT 27-SEP-2002
DEFINITION	Sequence 8 from Patent WO0226826.				
ACCESSION	AX503778				
VERSION	AX503778.1	GI:23385964			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1				
AUTHORS	Gerlach, V.L., Macdougall J.R., Smithson, G., Millet, I., Stone, D., Gunther, E., Ellerman, K., Grosse, W.M., Alsbrook, J.P., Lepley, D.M., Burgess, C.E., Padigaru, M., Kekuda, R., Spytek, K.A., Leach, M.D. and Shimkets, R.A.				
TITLE	Proteins and nucleic acids encoding same				
JOURNAL	Patent: WO 0226826-A 8 04-APR-2002;				
FEATURES	Curagen Corporation (US)				
source	Location/Qualifiers				
	1..6201				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
	Query Match	85.6%;	Score 5629,8;	DB 6;	Length 6201;
	Best Local Similarity	95.7%;	Pred. No. 0;		
	Matches 5929;	Conservative	0;	Mismatches	22;
				Indels	246;
				Gaps	4;
QY	19	ATGTTGAAGTTCAAATATGGAGCGCGAAATCCCTTTGGATGCTGGTGTGTGAACCCATT	78		
Db	1	ATGTTGAAGTTCAAATATGGAGCGCGAAATCCCTTTGGATGCTGGTGTGTGAACCCATT	60		
QY	79	GCCAGCGGGCTCCAGGCTGAATCTCTTCTCCAGGGGAAACACACCTTTATGACTCAA	138		
Db	61	GCCAGCGGGCTCCAGGCTGAATCTCTTCTCCAGGGGAAACACACCTTTATGACTCAA	120		
QY	139	CAGCAGATGTCCTCTTTCCGAGAAAGGATATAGATGCCCTCTTTGTTCTCTTTGAA	198		
Db	121	CAGCAGATGTCCTCTTTCCGAGAAAGGATATAGATGCCCTCTTTGTTCTCTTTGAA	180		
QY	199	GAATGCAGTCAGCTGCTGTGATGAAGATTAAACACGTGAGCACTTTGTCGGAGATAT	258		
Db	181	GAATGCAGTCAGCTGCTGTGATGAAGATTAAACACGTGAGCACTTTGTCGGAGATAT	240		
QY	259	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAAGTCAGA	318		
Db	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAAGTCAGA	300		
QY	319	AGTCTTGAGTTGTGGTCACTTTGCTGAGTCGAGTCGAGTCGAGCAAGCAACCGGG	378		
Db	301	AGTCTTGAGTTGTGGTCACTTTGCTGAGTCGAGTCGAGTCGAGCAAGCAACCGGG	360		
QY	379	GACATCTATGCTATGAAAGTGATGAAGAAAGGCTTTATTGGCCCGAGGAGCGTTTCA	438		
Db	361	GACATCTATGCTATGAAAGTGATGAAGAAAGGCTTTATTGGCCCGAGGAGCGTTTCA	420		
QY	439	TTTTTTGAGAAAGCGGACATATATCTCGAGGACCAAGCCGCTGGATCCCCCAATTA	498		
Db	421	TTTTTTGAGAAAGCGGACATATATCTCGAGGACCAAGCCGCTGGATCCCCCAATTA	480		
QY	499	CAGTATGCCCTTCAGGACAAAATACACCTTTATCTGATGGAGGAATATACGCTGGAGGG	558		
Db	481	CAGTATGCCCTTCAGGACAAAATACACCTTTATCTGATGGAGGAATATACGCTGGAGGG	540		
QY	559	GACTTGCTGTCACCTTTGATATAGATAGGACAGTTAGATGAACACCTGATACAGTTT	618		
Db	541	GACTTGCTGTCACCTTTGATATAGATAGGACAGTTAGATGAACACCTGATACAGTTT	600		
QY	619	TACCTAGCTCAGCTGATTTTGGCTGTTCCAGCGCTTCATCTGATGGGATACGTCATCGA	678		

Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTCTACAGGTTTCATCTGATGGGATACCTGTCATCGG 660
QY 679 GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 738
Db 661 GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
QY 739 GATCTGCCGCGGAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTTGGGACC 798
Db 721 GGATCTGCCGCGGAAAATGAATTCAAAACA ---GGTGAATGCCAAACTCCCGATTTGGGACC 777
QY 799 CCAGATTACATGCTCTCTGAAGTGCCTGACTGTGATCAACCGGGATGGAAAAGCCACCTAC 858
Db 778 CCAGATTACATGCTCTCTGAAGTGCCTGACTGTGATCAACCGGGATGGAAAAGCCACCTAC 837
QY 859 GGCTTGAATGTGATCTGGTGGTCACTGGGCGTGAATTCCTATGAGATGAATTTATGGGAGA 918
Db 838 GGCTTGAATGTGATCTGGTGGTCACTGGGCGTGAATTCCTATGAGATGAATTTATGGGAGA 897
QY 919 TCCGCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACATATATGAATTTCCAGCGG 978
Db 898 TCCGCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACATATATGAATTTCCAGCGG 957
QY 979 TTTTGGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1038
Db 958 TTTTGGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1017
QY 1039 TTGTTGTGGGCGCAGAAAGAGACTCAAGTTTGAAGGTTTGGTCCCATCTTCTTC 1098
Db 1018 TTGTTGTGGGCGCAGAAAGAGACTGAAATTTGAAGGTTTGGTCCCATCTTCTTC 1077
QY 1099 TCTAAAATTTGACTGGAAACAATTCCTCTCTCCGCCCTTCGTTCCCACTCAAG 1158
Db 1078 TCTAAAATTTGACTGGAAACAATTCCTCTCTCCGCCCTTCGTTCCCACTCAAG 1137
QY 1159 TCTGAGATGACACTTCCAAATTTTGATGAACACAGAGAAGATTCGTTGGGTTTCATCTCT 1218
Db 1138 TCTGAGATGACACTTCCAAATTTTGATGAACACAGAGAAGATTCGTTGGGTTTCATCTCT 1197
QY 1219 CCGTGGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1278
Db 1198 CCGTGGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1257
QY 1279 TACAGCAAGCCTGCGGATTTCTTGATGATCTGACTCTGTGTGTCGGGTTCGACTCC 1338
Db 1258 TACAGCAAGCCTGCGGATTTCTTGATGATCTGACTCTGTGTGTCGGGTTCGACTCC 1317
QY 1339 CCGTGGCAAGCTAGCTCCATGGBAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1398
Db 1318 CCGTGGCAAGCTAGCTCCATGGBAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1377
QY 1399 TCTCAGGACAAGTGTACAAAGATGGAGCAGGAATGACCCTGTTACATCGGAGGTGCA 1458
Db 1378 TCTCAGGACAAGTGTACAAAGATGGAGCAGGAATGACCCTGTTACATCGGAGGTGCA 1437
QY 1459 GAGGTGGAGGCTGTCTGTAGTCAAGAGGAGGTGGAGCTGAGCCCTCAGACTCAGAGA 1518
Db 1438 GAGGTGGAGGCTGTCTGTAGTCAAGAGGAGGTGGAGCTGAGCCCTCAGACTCAGAGA 1497
QY 1519 TCCCTCTCGAGCAGGACCTTGTACTCATCACAGAAATGCAAGTGTAAAGCGAAGT 1578
Db 1498 TCCCTCTCGAGCAGGACCTTGTACTCATCACAGAAATGCAAGTGTAAAGCGAAGT 1557
QY 1579 TTGGAGCAAGCAGGATGGAGGTGTCAGAGGAGGTGACAAAGCACTGAGCTTCCAT 1638
Db 1558 TTGGAGCAAGCAGGATGGAGGTGTCAGAGGAGGTGACAAAGCACTGAGCTTCCAT 1617
QY 1639 GATATCAGAGCAGAGCCGGAAGCTTCCAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1698
Db 1618 GATATCAGAGCAGAGCCGGAAGCTTCCAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1677
QY 1699 GTGGAAGAAATAGGTTGATGAATTCAGTTGGAAAGAGGATCTTGTCTCAGCAAGAGA 1758
Db 1678 GTGGAAGAAATAGGTTGATGAATTCAGTTGGAAAGAGGATCTTGTCTCAGCAAGAGA 1737

QY 1759 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAATTCAG 1818
Db 1738 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAATTCAG 1797
QY 1819 CGGAAACGCGACAGATCTCAGCATAACTCTTGAAGCTTAAGGATCAAGGGAACCTTGAA 1878
Db 1798 CGGAAACGCGACAGATCTCAGCATAACTCTTGAAGCTTAAGGATCAAGGGAACCTTGAA 1857
QY 1879 GTGGAGAAATATCGAAACTGGGAGAAATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938
Db 1858 GTGGAGAAATATCGAAACTGGGAGAAATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1917
QY 1939 CTCCAGAGAAATCTGGAGAAAGGCT----- 1962
Db 1918 CTCCAGAGAAATCTGGAGAAAGGCTGTAAGACCAGCAGCAGGAGGCCACCGAGCTGCTGCAG 1977
QY 1963 -----GCAAAAGGAGCGAGCGGAGAGGAGCTGGAGAAGCTGCAGAAACCGAGAG 2010
Db 1978 AATATCCGCGCAGCAAGAGGAGCGAGCGGAGAGCTGGAGAAGCTGCAGAAACCGAGAG 2037
QY 2011 GATTCTTCTGAAGCATCAGAAAGAGCTGGTGGAAAGCTGAGGAAACCGCCCATTTCTCTG 2070
Db 2038 GATTCTTCTGAAGCATCAGAAAGAGCTGGTGGAAAGCTGAGGAAACCGCCCATTTCTCTG 2097
QY 2071 GAGAAACAAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGAAAGGATGAC 2130
Db 2098 GAGAAACAAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGAAAGGATGAC 2157
QY 2131 ATCCAGACAAATTCCTAACAGATCCAGCAGATGGCTGATAAATTTCTGAGCTCGAGAG 2190
Db 2158 ATCCAGACAAATTCCTAACAGATCCAGCAGATGGCTGATAAATTTCTGAGCTCGAGAG 2217
QY 2191 AAACATCGGAGGCGCCAAAGTCTCAGCCAGACACTAGAACTGACCTGAAACAGAAAGAG 2250
Db 2218 AAACATCGGAGGCGCCAAAGTCTCAGCCAGACACTAGAACTGACCTGAAACAGAAAGAG 2277
QY 2251 CAGCACTATGAGGAAAGATTTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGGTGAC 2310
Db 2278 CAGCACTATGAGGAAAGATTTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGGTGAC 2337
QY 2311 AAGGAGACCTGAGAAACATGATCCAGAGACACAGAGGAGGCGCCCATGAGAGGCGCAA 2370
Db 2338 AAGGAGACCTGAGAAACATGATCCAGAGACACAGAGGAGGCGCCCATGAGAGGCGCAA 2397
QY 2371 ATTCTCAGCGAAACAGAAAGGCGGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2430
Db 2398 ATTCTCAGCGAAACAGAAAGGCGGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2457
QY 2431 CAGAGGATTTGGAACTGTCTGAAGCCAATAAATTTGCAGCAAAATAGCAGCTCTTTTACC 2490
Db 2458 CAGAGGATTTGGAACTGTCTGAAGCCAATAAATTTGCAGCAAAATAGCAGCTCTTTTACC 2517
QY 2491 CAAAGGAACTAAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550
Db 2518 CAAAGGAACTAAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577
QY 2551 CTGAGAGACAGGCTGGGAAATTTGGAGGCCAGAAACGAAAACTGGAGGAGCAGCTGGAG 2610
Db 2578 CTGAGAGACAGGCTGGGAAATTTGGAGGCCAGAAACGAAAACTGGAGGAGCAGCTGGAG 2637
QY 2611 AAGATCAGCCACCAAGACCCACAGTGACAGAACTCGGCTGTGGAATCTGGAGACAAAGATTG 2670
Db 2638 AAGATCAGCCACCAAGACCCACAGTGACAGAACTCGGCTGTGGAATCTGGAGACAAAGATTG 2697
QY 2671 CGGAGGTCTAGTCTAGAGACAGGAGCAGAAACTGGAGCTCAAGCGCAGCTCAGCAGAG 2730
Db 2698 CGGAGGTCTAGTCTAGAGACAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCAGCAGAG 2757
QY 2731 CTACAGCTCTCCCTGACGAGGCGCGAGTCAAGTTGACAGCCCTGACGGCTGCACGGCG 2790
Db 2758 CTACAGCTCTCCCTGACGAGGCGCGAGTCAAGTTGACAGCCCTGACGGCTGCACGGCG 2817

QY 2791 GCCCTGGAGCCAGCTTCGCCAGCGGAGACAGAGCTGGAGAGACCACAGCGAAGCT 2850
Db 2818 GCCCTGGAGCCAGCTTCGCCAGCGGAGACAGAGCTGGAGAGACCACAGCGAAGCT 2877
QY 2851 GAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910
Db 2878 GAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2937
QY 2911 CTTCTTAACAGTGTACTGTANTATCAGACCTGGAGGAGCAGCTTAACACAGGTGACCGAG 2970
Db 2938 CTTCTTAACAGTGTACTGTANTATCAGACCTGGAGGAGCAGCTTAACACAGGTGACCGAG 2997
QY 2971 GACAAACGCTGAACCTCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030
Db 2998 GACAAACGCTGAACCTCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3057
QY 3031 GCGCCAAACGACGAGATGTGTAACAATCGGAAGTGAAGTGGACCAATCTCCGCGGGAGATC 3090
Db 3058 GCGCCAAACGACGAGATGTGTAACAATCGGAAGTGAAGTGGACCAATCTCCGCGGGAGATC 3117
QY 3091 ACGGAACGAGAGATCGAGCTTACCAGCCAGAGCAACAGATGGAGGCTCTGAAGACCAAG 3150
Db 3118 ACGGAACGAGAGATCGAGCTTACCAGCCAGAGCAACAGATGGAGGCTCTGAAGACCAAG 3177
QY 3151 TGCACCATGCTGGAGGAACAGGTCTATGGATTGGAGGCCCTAAACAGATGAGCTGCTAGAA 3210
Db 3178 TGCACCATGCTGGAGGAACAGGTCTATGGATTGGAGGCCCTAAACAGATGAGCTGCTAGAA 3237
QY 3211 AAAGAGCGCAGTGGAGGCCCTGGAGGAGCGTCTCGGCTGATGAGAAATCCAGTTTCAG 3270
Db 3238 AAAGAGCGCAGTGGAGGCCCTGGAGGAGCGTCTCGGCTGATGAGAAATCCAGTTTCAG 3297
QY 3271 TGTCCGGTTCGAGAGCTGCAGAGAAATGTGGACACCGAGAAACAGAGAGGCGCAGAGGCC 3330
Db 3298 TGTCCGGTTCGAGAGCTGCAGAGGATGCTGGACACCGAGAAACAGAGAGGCGCAGAGGCC 3357
QY 3331 GATCAGCGATCACCAGTCTCCGAGGTGTGGAGCTGGCAGTGAAGAGGACACAGGCT 3390
Db 3358 GATCAGCGATCACCAGTCTCCGAGGTGTGGAGCTGGCAGTGAAGAGGACACAGGCT 3417
QY 3391 GAGATTCTCGCTCGCAGCAGCTCTCAAAGAGCAGAGCTTGAAGCGCAGAGCTCTCT 3450
Db 3418 GAGATTCTCGCTCGCAGCAGCTCTCAAAGAGCAGAGCTTGAAGCGCAGAGCTCTCT 3477
QY 3451 GACAAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3510
Db 3478 GACAAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3537
QY 3511 CAGCAGAGCTGGAGACTGAAAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3570
Db 3538 CAGCAGAGCTGGAGACTGAAAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3597
QY 3571 TTACAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTGCAA 3630
Db 3598 TTACAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTGCAA 3657
QY 3631 GAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTATCAGCTG 3690
Db 3658 GAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTATCAGCTG 3717
QY 3691 GAAACATTTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCCTCAA 3750
Db 3718 GAAACATTTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCCTCAA 3777
QY 3751 CAAAACAAACTCATTTGATTTTGTCAAGCAAAAATGGACCAACCTGCTTAAAGAAAAAG 3810
Db 3778 CAAAACAAACTCATTTGATTTTGTCAAGCAAAAATGGACCAACCTGCTTAAAGAAAAAG - - 3835
QY 3811 GGTTTATTAGTCGAGGAAAGAGGACCTGCTTTACCCACACAGGTTTCTCTGAGTAC 3870
Db 3836 -----AGTGCCTCTGAGTAC 3852
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGTGTGCAGACTGAGGAAGCC 3930

Db 3853 AATGAGCTGAAGCTTGGCCCTGGAGAGGAGAAAGCTTCGCTGTGCAGAGCTAGAGGAAGCC 3912
QY 3931 CTTTCAGAAAGACCCGCAATCGAGCTTCGGTTCGCGCCGGAGGAAGCTGCCACCGCAAGCA 3990
Db 3913 CTTTCAGAAAGACCCGCAATCGAGCTTCGGTTCGCGCCGGAGGAAGCTGCCACCGCAAGCA 3972
QY 3991 ACGGACCAACCAACCAATCCAGCCAGCCAGCCAGCCAGCGGAGGAGAGATCGCCATGTCGCG 4050
Db 3973 ACGGACCAACCAACCAATCCAGCCAGCCAGCCAGCCAGCGGAGGAGAGATCGCCATGTCGCG 4032
QY 4051 ATCGTGGGTCGCGCAGAGACCCAGCCAGTGCATGAGCCTGCTGGCCCGCCCATCCAGC 4110
Db 4033 ATCGTGGGTCGCGCAGAGACCCAGCCAGTGCATGAGCCTGCTGGCCCGCCCATCCAGC 4092
QY 4111 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTTCGGCGTCTTAAGBAACCGCATGCAC 4170
Db 4093 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTTCGGCGTCTTAAGBAACCGCATGCAC 4152
QY 4171 CACAATATTCTCAACCGATTCAACGTAGGACTGAAACATGCGAGCCACAAAGTGTGCTGTG 4230
Db 4153 CACAATATTCTCAACCGATTCAACGTAGGACTGAAACATGCGAGCCACAAAGTGTGCTGTG 4212
QY 4231 TGTCTGATACCGTGCATCTTTGGACGCGAGGATCCAAATGTCTCGAAATGTCAAGTGTG 4290
Db 4213 TGTCTGATACCGTGCATCTTTGGACGCGAGGATCCAAATGTCTAGAAATGTCAAGTGTG 4272
QY 4291 TGTCAACCCCAAGTCTCCAGTCTTCGACAGCCACCTGCGGCTTGCCTGTAATATGCC 4350
Db 4273 TGTCAACCCCAAGTCTCCAGTCTTCGACAGCCACCTGCGGCTTGCCTGTAATATGCC 4332
QY 4351 ACACACTTCAACCGAGGCTTCTCCGCTGACAAATGAACCTCCGAGTCTCCAGACCAAG 4410
Db 4333 ACACACTTCAACCGAGGCTTCTCCGCTGACAAATGAACCTCCGAGTCTCCAGACCAAG 4392
QY 4411 GAGCCAGCAGCAGCTTGCACCTGGAGGGTGGATGAAGTGCAGGAATACAAACGA 4470
Db 4393 GAGCCAGCAGCAGCTTGCACCTGGAGGGTGGATGAAGTGCAGGAATACAAACGA 4452
QY 4471 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCATTTAT 4530
Db 4453 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCATTTAT 4512
QY 4531 GACAAATGAAGCCAGAGAGCTGGACAGAGCCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4590
Db 4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4572
QY 4591 GACGGGATGTATCTATTCTATGCTGCGTTCGCTTCCGAACTCGCAATACAGCCAAA 4650
Db 4573 GACGGGATGTATCTATTCTATGCTGCGTTCGCTTCCGAACTCGCAATACAGCCAAA 4632
QY 4651 GCA----- 4653
Db 4633 GCAGATGTCCATACATACACTGAAGTGAATCTCACCCGCACACCACTGCTGCGCCCGG 4692
QY 4654 ----- 4653
Db 4693 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCGCTTA 4752
QY 4654 -----GAAAAAGCAGAAGCTGATAAATG 4680
Db 4753 GAATCAGTTCGAGGCTGGAGAGTTTCTAGGGAAGAAAGCAGAGCTGATCTAAATG 4812
QY 4681 CTTGGAAATCTCCCTGCTGAAACTGGAAGTGAACCTGTAGACATGAATGACAGCTG 4740
Db 4813 CTTGGAAATCTCCCTGCTGAAACTGGAAGTGAACCTGTAGACATGAATGACAGCTG 4872
QY 4741 CCCTTCAGTACAGGTTGTTGTTGGGCAACCGAGGAAGGCTCTACGCCCTGAATGTC 4800
Db 4873 CCCTTCAGTACAGGTTGTTGTTGGGCAACCGAGGAAGGCTCTACGCCCTGAATGTC 4932
QY 4801 TTGAAAACTCCCTAAACCCATGTCAGGAATTTGGAGCAGTCTTCCAAATTTATATTC 4860

4933 TTGAAAACTCCCTAACCCATGTCCAGGAAATGGAGCAGTCTTCCAAATTTATATTATC 4992
4861 AAGGACCTGGAGAGCTACTCATGTAGCAGGAGAAAGCGGCACTGTGTCTGTGGAC 4920
4993 AAGGACCTGGAGAGCTACTCATGTAGCAGGAGAAAGCGGCACTGTGTCTGTGGAC 5052
4921 GTGAGAAAGTGAAGACAGTCCCTGGCCAGTCCACCTGCTGCGCCAGCCGACATCTCA 4980
5053 GTGAGAAAGTGAAGACAGTCCCTGGCCAGTCCACCTGCTGCGCCAGCCGACATCTCA 5112
4981 CCCAACATTTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5040
5113 CCCAACATTTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5172
5041 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAGTCTGATCTTCCGCTACCAACGAAAC 5100
5173 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAGTCTGATCTTCCGCTACCAACGAAAC 5232
5101 CTCAGCAATATCTGCATCTCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160
5233 CTCAGCAATATCTGCATCTCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5292
5161 TTCACCAATATAGATATCTTCATTTGAAACCAATAAATTTCTACGAATCGACATGAAGCAG 5220
5293 TTCACCAATATAGATATCTTCATTTGAAACCAATAAATTTCTACGAATCGACATGAAGCAG 5352
5221 TACACGCTCAGGAAATCTCTGGATAAGATGACCAATCTCTGGCACCTGCTGTGTTGCC 5280
5353 TACACGCTCAGGAAATCTCTGGATAAGATGACCAATCTCTGGCACCTGCTGTGTTGCC 5412
5281 GCCTCTTCCAAAGCTTCCCTGCTCAATCTGTCAGGTGAACAGCGCAGGCGAGGAG 5340
5413 GCCTCTTCCAAAGCTTCCCTGCTCAATCTGTCAGGTGAACAGCGCAGGCGAGGAG 5472
5341 GAGTACTTGTGTTTCCAGCAATTTGGAGTCTTCTGGATTTCTTACGGAAGAGCTAGC 5400
5473 GAGTACTTGTGTTTCCAGCAATTTGGAGTCTTCTGGATTTCTTACGGAAGAGCTAGC 5532
5401 CGCACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5460
5533 CGCACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5592
5461 CTGTTTGTGACCCACTTCACTCACTGGAAGTATTTGAGATCCAGGACGCTCCTCAGCA 5520
5593 CTGTTTGTGACCCACTTCACTCACTGGAAGTATTTGAGATCCAGGACGCTCCTCAGCA 5652
5521 GGGACCCCTGCCGAGGCTACCTGGACATCCCGAAACCCCGCGTACCTGGGCGCTGCCATT 5580
5653 GGGACCCCTGCCGAGGCTACCTGGACATCCCGAAACCCCGCGTACCTGGGCGCTGCCATT 5712
5581 TCCTCAGGACGATTTACTTGGGCTCTCATACAGGATAAATTAAGGTCAATTTGCTGC 5640
5713 TCCTCAGGACGATTTACTTGGGCTCTCATACAGGATAAATTAAGGTCAATTTGCTGC 5772
5641 AAGGGAAACCTCGTGAAGAGTCCGACCTGAAACACACCGGGGCGCTGCCCTCCGC 5700
5773 AAGGGAAACCTCGTGAAGAGTCCGACCTGAAACACACCGGGGCGCTGCCCTCCGC 5832
5701 AGCAGCCCAACAGCGAGGCGCCACCGATCAACAGGACATCAACGAAGCGGTGGCC 5760
5833 AGCAGCCCAACAGCGAGGCGCCACCGATCAACAGGACATCAACGAAGCGGTGGCC 5892
5761 TCCAGCCCGCGCGCCGAGGCGCCAGCACCAGCGAGAGCGCAAGCACACCCACCGC 5820
5893 TCCAGCCCGCGCGCCGAGGCGCCAGCACCAGCGAGAGCGCAAGCACACCCACCGC 5952
5821 TACCGCGAGGGCGGACCGAGTCTGCGCAGGACAACTCTCTGGCGCCCTCTGAGGAG 5880
5953 TACCGCGAGGGCGGACCGAGTCTGCGCAGGACAACTCTCTGGCGCCCTCTGAGGAG 6012
5881 GAGAGTCTCCCGCGGAGTCTCAGCAGCGGAGAGCGGTCCCCCGGAGGCTGTTT 5940
6013 GAGAGTCTCCCGCGGAGTCTCAGCAGCGGAGAGCGGTCCCCCGGAGGCTGTTT 6072

QY 5941 GAAGACAGCAGCGGCGCGCTGCTGCGGAGCCGCTGAGGACCCCGCTGTCCCGAGTG 6000
Db 6073 GAAGACAGCAGCAGCGGCGCGCTGCTGCGGAGCCGCTGAGGACCCCGCTGTCCCGAGTG 6132
QY 6001 AACCAAGGGAAGAGGCA 6017
Db 6133 AACAGGTGAGGAGCA 6149
RESULT 11
LOCUS AY257469 8576 bp mRNA linear PRI 23-APR-2003
DEFINITION Homo sapiens rho/rac-interacting citron kinase (CIT) mRNA, complete cds.
ACCESSION AY257469
VERSION AY257469.1 GI:30088969
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 8576)
Huang, C.Q., Wu, S.L., Shan, Y.X., Liu, S. and Xiao, P.J.
Direct Submission
Submitted (18-MAR-2003) Department of Biochemistry and Molecular
Biology, Basic Medicine, Suzhou University, Renmin Road 48,
Suzhou, Jiangsu 215007, China
Location/Qualifiers
1..8576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24.23"
1..8576
/gene="CIT"
/note="synonym: CRIK"
54..6137
/gene="Citr"
/note="serine/threonine protein kinases"
/codon_start=1
/product="rho/rac-interacting citron kinase"
/protein_id="AAPI3528.1"
/translation="MLKFKYGNPLDAGAAEPIASASRLNLFPOGKPPFTQQMS
PLREGIDALFVLPFECQPALMKIKNVFNFKYSDTIAELQELQPSAKDPEVRS
VGGHFAEVQVREKATGDIYAMKMKKALLAQEQSFEEBRLSRSTSPWIPQL
OYAFQDNHUYLVMEYQPGDLSLLNRYEDOLDENLIQYLAELIYAVHVSVMGV
HRDIPENILVDRTGHIKLVDFGSAAMNSKVNNAKLPIGTPDYNAVELTVMGDD
KTYGLDCDWSVGVITAYEMIVGRSPASGTSARTFNNIMNFORFLKFPDPPKVSDF
LDLIQSLCCQKRLAFEGLCCHPFPFSKIDWNINRSPFPFVTLSDDDTSNDFSE
KNWSVSCQSPSPSGSEELPFVGSYKALGILGRSESVSGSDSPAKTSMMKK
LLKSKELQSDQCHKMEQMTDLHRRVSEVAVLSQKEVELKATEQSRSLLEQDLA
TYTSCSLKRSLEQARMEVQSDKALQLLHDIRQSKLOIKQEQVQAEVRL
MNQLEEDLVARRRDLVSELSRSLAAEFKFKATECOHKLKAKDCKGPEVEY
AKKLEINAEQOLKIQLQELKAVKASTELLQIRQAKERAERLEIKLONRDS
SEGIRKLVAEELSEKHEAOVSAOHLEVLHAKQEQVVEEIKVLNDQIKKLAQKE
TLENMQRHEEAHEKQKILSEOKAMINAMDSKIRLEQRIVELSEANKLIANSIPT
QRNKAQEMISBLRQOKFYLEQAKLEAQNKRLEQLEKI SHQHSQKRLLELET
RLRSELEHEQKLEQLQTLQSLQERESQTLQALQARALESQKRAQTELEBET
TAAEBEIQALHARDEIQKFDALRNSCTVTDLQNLQEDNEDNAELNKNFYLSK
QLDEASANDIQLRASEVDHLREITEREMQTSQKTEALKTCTMLEEQMDLE
ALNDELEKROEAWRSVLGDEKQFECVRLEQMLDTEKOSRARADQRTESQV
VELAVKEKAEILALQALKEQKKAESLSDKNDLKKHAMLENARSLOQKLETER
ELQKRLLEQAKLQCCQMDLQKNIHIFLTQGLQELADLLKTERSDLYQLENIQVL
YSHKVMGEGTISQTKLIDLOAKMDQPAKKKKGLFSRRKEDPALPTQVLPQYINELK
LALKEKARCAREALEALQKTRIELRSEAAHRRKATDHPHPTPATATQQAIAISAV
RSEPHQFSAMSLAPPSRRKESSTPEFSRRLKERMHNIHPRFVGNLNRKCAV
CLDTVHFQROAKCLQVNMCHPKCSTCLPATCGLPAEYATHTFEACFCRDKNSPGLQ
TKPSSSLHLEGMKYPNNRKNQCGWDRKIIVLEGSKVLIYDNEAREAGQVVEFFE
LCPLDGDVSIHGAVGASELANAKADVPYILKMHESHPTTCWPGRTLYLLAPSPKQ
RWTALSVVAGGRVREKAEADAKLLGNSLLKLEGGDRLDMMNCTLPFSQVQVVLGTE

QY 1981 AGGAGGCTGGAGAGCTGCGAGAACCGAGAGGATTCCTTCTGAAGGCATCAGAAAGAGCTG 2040
Db |||||
2015 --GCCACCAGCTGTGTCAGAAATATCCGCCAGGCAAGAGGAGCGAGCCGAGAGGAGCTG 2072
QY |||||
2041 GTGGAGCTGAGAGACGCCGCCATCTCTGGAGAACAGAGTAAAGAGATAGAGACCATG 2100
Db |||||
2073 GAGAAGCTGCAGAACCGAGAGGATTCCTTCTGAAGGCATCAGAAAGAGCTGTGGTGG----- 2127
QY |||||
2101 GAGCGTAGAGAAACAGACTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAG 2160
Db ----- 2127
2161 ATGGCTGATAAATTTCTGGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGGCCAG 2220
Db |||||
2128 -----AAGCTGAGGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGGCCAG 2177
QY |||||
2221 CAGCTAGAGTGCACCTGAAACAGAAAGAGCAGCAGTATGAGGAAAAGATTAAAGTGTG 2280
Db |||||
2178 CACCTAGAGTGCACCTGAAACAGAAAGAGCAGCAGTATGAGGAAAAGATTAAAGTGTG 2237
QY |||||
2281 GACAACTCAGATAAAGAAAGCTGGCTGACAGAGGAGACACTGAGAGAAACATGATGCGAGAG 2340
Db |||||
2238 GACAATCAGATAAAGAAAGCCTGGCTGACAGGAGACACTGAGAGAACATGATGCGAGAG 2297
QY |||||
2341 CAGCAGAGAGGCCCATGAGAAAGGCAAAATTTCTCAGCGCAACAGAAAGCGCATGATCAAT 2400
Db |||||
2298 CACGAGAGAGGCCCATGAGAAAGGCAAAATTTCTCAGCGCAACAGAAAGCGCATGATCAAT 2357
QY |||||
2401 GCTATGATTCAGAGTACAGATCCCTGGAAACAGAGAGATTGTGGAATGTCTGAAGCCAAAT 2460
Db |||||
2358 GCTATGATTCAGAGTACAGATCCCTGGAAACAGAGAGATTGTGGAATGTCTGAAGCCAAAT 2417
QY |||||
2461 AACTTGCACCAATAGCAGCTCTTTTACCCAAAGGAACATGAAAGGCCCAAGAGAGATG 2520
Db |||||
2418 AACTTGCACCAATAGCAGCTCTTTTACCCAAAGGAACATGAAAGGCCCAAGAGAGATG 2477
QY |||||
2521 ATTCTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGCTCGGGAAGTTGGAGGCC 2580
Db |||||
2478 ATTCTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGCTCGGGAAGTTGGAGGCC 2537
QY |||||
2581 CAGAACCGAAAATGGAGGAGAGCTGGAGAGATCAGCCACAGACACACATGACAG 2640
Db |||||
2538 CAGAACCGAAAATGGAGGAGAGCTGGAGAGATCAGCCACAGACACACATGACAG 2597
QY |||||
2641 AATCGGCTGCTGGAACCTGGAGACAAGATTGCGGGAGGTGAGTCTAGAGCACGAGAGCAG 2700
Db |||||
2598 AATCGGCTGCTGGAACCTGGAGACAAGATTGCGGGAGGTGAGTCTAGAGCACGAGAGCAG 2657
QY |||||
2701 AACTGGAGCTCAAGGCCAGCTCAGAGAGTACAGTCTCTCCCTGAGAGGCGCGAGTCA 2760
Db |||||
2658 AACTGGAGCTCAAGGCCAGCTCAGAGAGTACAGTCTCTCCCTGAGAGGCGCGAGTCA 2717
QY |||||
2761 CAGTTGACAGCCCTGAGGCTGACAGGCGGCCCTGGAGAGCCAGTTCGCCAGCGAG 2820
Db |||||
2718 CAGTTGACAGCCCTGAGGCTGACAGGCGGCCCTGGAGAGCCAGTTCGCCAGCGAG 2777
QY |||||
2821 ACAGAGCTGGAAGAGACCAAGAGAGTGAAGAGGAGATCCAGCAGTCTCAGGACAT 2880
Db |||||
2778 ACAGAGCTGGAAGAGACCAAGAGAGTGAAGAGGAGATCCAGCAGTCTCAGGACAT 2837
QY |||||
2881 AGAGATGAATTCAGCGCAATTTGATGCTCTTCTGTAACAGCTGATCTGATATCAGAG 2940
Db |||||
2838 AGAGATGAATTCAGCGCAATTTGATGCTCTTCTGTAACAGCTGATCTGATATCAGAG 2897
QY |||||
2941 CTGGAGAGCAGCTAAACAGCTGACCGGAGCAACCGCTGAATCAACCAACAAAATTTC 3000
Db |||||
2898 CTGGAGAGCAGCTAAACAGCTGACCGGAGCAACCGCTGAATCAACCAACAAAATTTC 2957
QY |||||
3001 TACTTGTCCAAACAACTCGATGAGGCTCTGGCGCCCAACGAGAGATTGTACAACTGCGA 3060
Db |||||
2958 TACTTGTCCAAACAACTCGATGAGGCTCTGGCGCCCAACGAGAGATTGTACAACTGCGA 3017
QY |||||
3061 AGTGAAGTGGACCATCTCCGCCGGAGATCAGGAAACGAGAGATGCGAGCTTACCAGCCAG 3120

Db |||||
3018 AGTGAAGTGGACCATCTCCGCCGGAGATCAGGAAACGAGAGATGACAGCTTACCAGCCAG 3077
QY |||||
3121 AAGCAAAACGATGGAGGCTCTGAAGACCAACGTCACCATGTCTGGAGGAACAGGTTCATGGAT 3180
Db |||||
3078 AAGCAAAACGATGGAGGCTCTGAAGACCAACGTCACCATGTCTGGAGGAACAGGTTCATGGAT 3137
QY |||||
3181 TTGGAGGCCCTTAAACGATAGCTCTGTAAGAAAGAGCGGAGTGGAGGCTTGGAGAGC 3240
Db |||||
3138 TTGGAGGCCCTTAAACGATAGCTCTGTAAGAAAGAGCGGAGTGGAGGCTTGGAGAGC 3197
QY |||||
3241 GTCTTGGGTGATGAGAAATCCAGTTTGTAGTGTCCGGTTTCGAGAGCTGACAGAAATCCTG 3300
Db |||||
3198 GTCTTGGGTGATGAGAAATCCAGTTTGTAGTGTCCGGTTTCGAGAGCTGACAGAAATGCTG 3257
QY |||||
3301 GACACCGAGAAACAGAGCAGGCGGAGCCGATCAGCGGATCACCGAGTCTCGCCAGGTG 3360
Db |||||
3258 GACACCGAGAAACAGAGCAGGCGGAGCCGATCAGCGGATCACCGAGTCTCGCCAGGTG 3317
QY |||||
3361 GTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCTCGCTCTGACAGAGGCTCTCAAA 3420
Db |||||
3318 GTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCTCGCTCTGACAGAGGCTCTCAAA 3377
QY |||||
3421 GAGCAGAGCTGAAGGCGGAGAGCCTCTCTGCAAGCTCAATGACCTGGAGAAAGAGCAT 3480
Db |||||
3378 GAGCAGAGCTGAAGGCGGAGAGCCTCTCTGCAAGCTCAATGACCTGGAGAAAGAGCAT 3437
QY |||||
3481 GCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGAAACGAGGCTC 3540
Db |||||
3438 GCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGAAACGAGGCTC 3497
QY |||||
3541 AAACAGAGGCTTCTGGAGAGCAGCAAGCCAAATTCAGCAGCAGATGGACCTGCGAAGAAAT 3600
Db |||||
3498 AAACAGAGGCTTCTGGAGAGCAGCAAGCCAAATTCAGCAGCAGATGGACCTGCGAAGAAAT 3557
QY |||||
3601 CACATTTTCCGCTCTGACTCAAGGAGCTCAAGAGAGCTCTAGATCGGGCTGATCTACTGAAG 3660
Db |||||
3558 CACATTTTCCGCTCTGACTCAAGGAGCTCAAGAGAGCTCTAGATCGGGCTGATCTACTGAAG 3617
QY |||||
3661 ACAGAAAGAGTGAATTTGGAGTATCAGCTGGAACAAATTCAGGTTCTCTATTCTCATGAA 3720
Db |||||
3618 ACAGAAAGAGTGAATTTGGAGTATCAGCTGGAACAAATTCAGGTTCTCTATTCTCATGAA 3677
QY |||||
3721 AAGTGAATAATGGAGGCACTATTTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCC 3780
Db |||||
3678 AAGTGAATAATGGAGGCACTATTTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCC 3737
QY |||||
3781 AAAATGGACCAACCTGCTAAAAAGAAAAAGGTTTATTTAGTCGACGGAAGAGGACCCCT 3840
Db |||||
3738 AAAATGGACCAACCTGCTAAAAAGAAAAAGGTTTATTTAGTCGACGGAAGAGGACCCCT 3797
QY |||||
3841 GCTTTACCCACACAGGTTCTCTGCACTACAATGAGCTGAAGCTGGCCCTGGAGAAAGGAG 3900
Db |||||
3798 GCTTTACCCACACAGGTTCTCTGCACTACAATGAGCTGAAGCTGGCCCTGGAGAAAGGAG 3857
QY |||||
3901 AAAGCTCGCTGTCAGAGCTAGAGAGCCCTTCAGAGAACCCGCACTCGAGCTCCGGTCC 3960
Db |||||
3858 AAAGCTCGCTGTCAGAGCTAGAGAGCCCTTCAGAGAACCCGCACTCGAGCTCCGGTCC 3917
QY |||||
3961 GCCCGGAGGAAGCTGCCACCCGCAAGCAACCCGCAACCCATCCAGCCAGCC 4020
Db |||||
3918 GCCCGGAGGAAGCTGCCACCCGCAAGCAACCCGCAACCCATCCAGCCAGCC 3977
QY |||||
4021 ACCCGGAGGAGAGAGATCGCATGTCGCCATGTCGGGTGCGCAGAGCAACCGAGTCC 4080
Db |||||
3978 ACCCGGAGGAGAGATCGCATGTCGCCATGTCGGGTGCGCAGAGCAACCGAGTCC 4037
QY |||||
4081 GCCATGAGCTGCTGGGCCCGCCATCCAGCGGAGAAAGAGTCTTCAATCCAGAGGAA 4140
Db |||||
4038 GCCATGAGCTGCTGGGCCCGCCATCCAGCGGAGAAAGAGTCTTCAATCCAGAGGAA 4097
QY |||||
4141 TTTAGTGGCGCTTTAAGGAAACGATGCAACAAATTTCTCACCGATTCAACGTAGGA 4200

Db 4098 TTTAGTCGGGCTCTTAAGGAACGATGACCAACAATATTCCTCACCGATTCAACGTAGGA 4157
Qy 4201 CTGAACATGCGAGCCACAAAGTGTCTGTGTCTGATACCGTGCACTTTTGGACGCCAG 4260
Db 4158 CTGAACATGCGAGCCACAAAGTGTCTGTGTCTGATACCGTGCACTTTTGGACGCCAG 4217
Qy 4261 GCATCCAAATGTCTCGAATGTCAAGTGATGTGTACCCCAAGTGTCTCAAGTGTCTGCGCA 4320
Db 4218 GCATCCAAATGTCTCGAATGTCAAGTGATGTGTACCCCAAGTGTCTCAAGTGTCTGCGCA 4277
Qy 4321 GCCACTCGGGCTTCCCTGCTGAATATGCCACACACTTCCAGGAGCCCTTCTGCCGTGAC 4380
Db 4278 GCCACTCGGGCTTCCCTGCTGAATATGCCACACACTTCCAGGAGCCCTTCTGCCGTGAC 4337
Qy 4381 AAAATGAATCTCCAGGCTTCCAGACCAAGGAGCCAGACGAGCTTGCACCTGGAAGGG 4440
Db 4338 AAAATGAATCTCCAGGCTTCCAGACCAAGGAGCCAGACGAGCTTGCACCTGGAAGGG 4397
Qy 4441 TGGATGAAGGTCGCCAGGAATAACAAAGAGACAGCAAGCTGGACAGGAAGTACATT 4500
Db 4398 TGGATGAAGGTCGCCAGGAATAACAAAGAGACAGCAAGCTGGACAGGAAGTACATT 4457
Qy 4501 GTCTCGGAGGATCAAAAGTCTCTATTATGACAAATGAAGCCAGAGAGCTGGACAGAGG 4560
Db 4458 GTCTCGGAGGATCAAAAGTCTCTATTATGACAAATGAAGCCAGAGAGCTGGACAGAGG 4517
Qy 4561 CCGTGGAGAAATTTGAGCTGTGCTTCCGACGGGATGTATCTATTATCATGTGCGCTT 4620
Db 4518 CCGTGGAGAAATTTGAGCTGTGCTTCCGACGGGATGTATCTATTATCATGTGCGCTT 4577
Qy 4621 GGTGCTTCGAACTCGCAATACAGCAAGCA----- 4653
Db 4578 GGTGCTTCGAACTCGCAATACAGCAAGCAAGATGCCATACATACTGAGATGGAA 4637
Qy 4654 ----- 4653
Db 4638 TCTCACCGCACACCACTGTCTGGCCGGGAAACCTCTACTTGTAGCTCCACGCTTC 4697
Qy 4654 ----- 4653
Db 4698 CTGCAAAACAGCGTGGGTCTACCGCTTAGAATCAGTTGTGCGAGGTGGAGATTCTT 4757
Qy 4654 ---GAAAAAGCAGAGCTGATCTAACTGCTTTGGAACCTCCCTGCTGAAACTGGAAGGT 4710
Db 4758 AGGAAAAAGCAGAGCTGATCTAACTGCTTTGGAACCTCCCTGCTGAAACTGGAAGGT 4817
Qy 4711 GATGACCGTCTAGACATGAACTGCAAGCTGCGCTTCACTGACAGGTGGTGTGTGGGC 4770
Db 4818 GATGACCGTCTAGACATGAACTGCAAGCTGCGCTTCACTGACAGGTGGTGTGTGGGC 4877
Qy 4771 ACCGAGGAAGGCTCTACCGCTGAAATGCTTTGAAAAAATCCCTTAACCCATGTCACAGGA 4830
Db 4878 ACCGAGGAAGGCTCTACCGCTGAAATGCTTTGAAAAAATCCCTTAACCCATGTCACAGGA 4937
Qy 4831 ATTGGAGAGCTCTCCAAATTTATATTATCAAGGACCTGGAGAGCTACTCATGTATAGCA 4890
Db 4938 ATTGGAGAGCTCTCCAAATTTATATTATCAAGGACCTGGAGAGCTACTCATGTATAGCA 4997
Qy 4891 GGAGAGCGCGGCACTGTGCTTTGGACGTGAAGAGTGAACAGTCCCTGCGCCAG 4950
Db 4998 GGAGAGCGCGGCACTGTGCTTTGGACGTGAAGAGTGAACAGTCCCTGCGCCAG 5057
Qy 4951 TCCACCTGCTGCCAGCCGACATCTCACCCAAACATTTTGAAGCTGTCAAGGGCTGC 5010
Db 5058 TCCACCTGCTGCCAGCCGACATCTCACCCAAACATTTTGAAGCTGTCAAGGGCTGC 5117
Qy 5011 CACTTTTGGGCGAGCAAGATTGGAACGGGCTCTGCATCTGTGCACCAATGCCCCAGC 5070
Db 5118 CACTTTTGGGCGAGCAAGATTGGAACGGGCTCTGCATCTGTGCACCAATGCCCCAGC 5177
Qy 5071 AAAGTCGTCTATTCTCGCTACACAGAAACCTCAGCAATATCTGATCCGGAAGAGATA 5130
Db 5178 AAAGTCGTCTATTCTCGCTACACAGAAACCTCAGCAATATCTGATCCGGAAGAGATA 5237

Qy 5131 GAGACCTCAGAGCCCTGCAGCTGTATCCACTTCAACAAATTAAGTATCTCATTTGGAACC 5190
Db 5238 GAGACCTCAGAGCCCTGCAGCTGTATCCACTTCAACAAATTAAGTATCTCATTTGGAACC 5297
Qy 5191 AATAAATTTCTACCAATCGACATGAAGCAGTACACCTCGAGGAATTCCTTGGATAGAAT 5250
Db 5298 AATAAATTTCTACCAATCGACATGAAGCAGTACACCTCGAGGAATTCCTTGGATAGAAT 5357
Qy 5251 GACCAATTCCTTGGCACTGTCTGTTTGGCGCTCTTCCAAACAGCTTCCCTGTCTCAATC 5310
Db 5358 GACCAATTCCTTGGCACTGTCTGTTTGGCGCTCTTCCAAACAGCTTCCCTGTCTCAATC 5417
Qy 5311 GTGAGGTGAACAGCCAGGCGAGGAGGAGTACTTGTCTGTGTTTCCACGAATTTGGA 5370
Db 5418 GTGAGGTGAACAGGCGAGGCGAGGAGGAGTACTTGTCTGTGTTTCCACGAATTTGGA 5477
Qy 5371 GTGTTCTGTGATTTCTTACGGAAGACGTAGCCGCACAGACGATCTCAAGTGGAGTCCCTTA 5430
Db 5478 GTGTTCTGTGATTTCTTACGGAAGACGTAGCCGCACAGACGATCTCAAGTGGAGTCCCTTA 5537
Qy 5431 CTTTGGCCTTTGCTTACAGAGAACCTTCTCTGTTTGTGACCCACTTCACTCACTGGAA 5490
Db 5538 CTTTGGCCTTTGCTTACAGAGAACCTTCTCTGTTTGTGACCCACTTCACTCACTGGAA 5597
Qy 5491 GTAATTGAGATCAGGACACGCTCTCAGCAGGACCCCTGCCGAGCGTACCTGGACATC 5550
Db 5598 GTAATTGAGATCAGGACACGCTCTCAGCAGGACCCCTGCCGAGCGTACCTGGACATC 5657
Qy 5551 CCGAACCCGCGCTTACCTTGGCCCTTGCAATTTCTCAGGAGCGATTTACTTGGGCTCTCA 5610
Db 5658 CCGAACCCGCGCTTACCTTGGCCCTTGCAATTTCTCAGGAGCGATTTACTTGGGCTCTCA 5717
Qy 5611 TACAGATGAATTAAGGTCATTTCTGCAAGGGAACCTCTGTAAGGAGTCCGGCACT 5670
Db 5718 TACAGATGAATTAAGGTCATTTCTGCAAGGGAACCTCTGTAAGGAGTCCGGCACT 5777
Qy 5671 GAAACACACCGGGCCCGTCTCAGCTCCGCGAGCAGCCCAACAGCGAGGCCCAACCCACG 5730
Db 5778 GAAACACACCGGGCCCGTCTCAGCTCCGCGAGCAGCCCAACAGCGAGGCCCAACCCACG 5837
Qy 5731 TACACAGACATCACAAGCGGTGGCTTCAGCCAGCGCGCCGCGAGGCGCCAGC 5790
Db 5838 TACAACAGACATCACAAGCGGTGGCTTCAGCCAGCGCGCCGCGAGGCGCCAGC 5897
Qy 5791 CACCCGAGAGCAACACACCCCTACCGGAGGCGGCGAGCCAGCTGCGCAGG 5850
Db 5898 CACCCGAGAGCAACACACCCCTACCGGAGGCGGCGAGCCAGCTGCGCAGG 5957
Qy 5851 GACAAGTCTCTGCGCCCGCTGAGCAGAGAAAGTCCCGCGCGCGGATGCTCAGCAG 5910
Db 5958 GACAAGTCTCTGCGCCCGCTGAGCAGAGAAAGTCCCGCGCGCGGATGCTCAGCAG 6017
Qy 5911 CCGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACAGCAGAGGCGCGGCTGCGCTGCG 5970
Db 6018 CCGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACAGCAGAGGCGCGGCTGCGCTGCG 6077
Qy 5971 GGAGCCGTGAGGACCCCGCTGTCCAGGTGAACAAGGGAAGGCGAGGTGCTCTCAA 6030
Db 6078 GGAGCCGTGAGGACCCCGCTGTCCAGGTGAACAAGGTCTGGGACCACTCTTCAATATA 6137
Qy 6031 GTTTTACGGTTAATCTGTACCTATTAT 6060
Db 6138 ATCTCAGCCAGAAAAACCAACTCTCTCATCT 6167

RESULT 12

AX574427 LOCUS AX574427 5877 bp DNA linear PAT 07-JAN-2003

DEFINITION Sequence 3 from Patent WO02059325.

ACCESSION AX574427

VERSION AX574427.1 GI:27551752

KEYWORDS

4051 ATGTCGCGTCCGACAGACACAGCCAGTGCATGACCTGCTGGCCCGCCATCCAGC 4110
Db |
4036 ATGTCGCGTCCGACAGACACAGCCAGTGCATGACCTGCTGGCCCGCCATCCAGC 4095
Qy |
4111 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCGCGGCTTTAAGGAAGCATGAC 4170
Db |
4096 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCGCGGCTTTAAGGAAGCATGAC 4155
Qy |
4171 CACAATATTCTCACCAGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTCTGTG 4230
Db |
4156 CACAATATTCTCACCAGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTCTGTG 4215
Qy |
4231 TGTCTGATACCGTGACCTTTGGACCCAGGATCCAAATGCTCGAATGTCAGTGATG 4290
Db |
4216 TGTCTGATACCGTGACCTTTGGACCCAGGATCCAAATGCTCGAATGTCAGTGATG 4275
Qy |
4291 TGTCAACCCAAAGTCTCCAGCTGCTTGCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4350
Db |
4276 TGTCAACCCAAAGTCTCCAGCTGCTTGCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4335
Qy |
4351 ACACATTCACCGAGGCTTTCGCTGTGACAAAATGAATCTCCAGGCTTCAGACCAAG 4410
Db |
4336 ACACATTCACCGAGGCTTTCGCTGTGACAAAATGAATCTCCAGGCTTCAGACCAAG 4395
Qy |
4411 GAGCCAGACAGAGCTTGACCTGGAGGCTGATGAAGTGCCTGAGGAATTAACAAACGA 4470
Db |
4396 GAGCCAGACAGAGCTTGACCTGGAGGCTGATGAAGTGCCTGAGGAATTAACAAACGA 4455
Qy |
4471 GGAACAGAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4530
Db |
4456 GGAACAGAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4515
Qy |
4531 GACAAATGAGCCAGAGAGCTGACAGAGCGGCTGAGAGAAATTTGAGCTGCGCTTCC 4590
Db |
4516 GACAAATGAGCCAGAGAGCTGACAGAGCGGCTGAGAGAAATTTGAGCTGCGCTTCC 4575
Qy |
4591 GAGGGGATGTATCTATTATGCTGCGTGGTTCGAACTCGCAATACAGCCAAA 4650
Db |
4576 GAGGGGATGTATCTATTATGCTGCGTGGTTCGAACTCGCAATACAGCCAAA 4635
Qy |
4651 GGA----- 4653
Db |
4636 GCAGATGTCCCATACATACCTGGAATCTACCCCGCACACACCTGCTGGCCCGGG 4695
Qy |
4654 ----- 4653
Db |
4696 AGAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGCTCACCCTTA 4755
Qy |
4654 -----GAAAAAGCAGAGCTGATGCTTAACTG 4680
Db |
4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAGAGCAGAGCTGATCTAACTG 4815
Qy |
4681 CTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTAGACATGAACTGCAAGCTG 4740
Db |
4816 CTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTAGACATGAACTGCAAGCTG 4875
Qy |
4741 CCCTTCAGTGACAGGTGTTGTTGGTGGGACCGAGGAAGGCTCTACCCCTGAAATGTC 4800
Db |
4876 CCCTTCAGTGACAGGTGTTGTTGGTGGGACCGAGGAAGGCTCTACCCCTGAAATGTC 4935
Qy |
4801 TTGAAAACCTCCCTAACCTATGCTCCAGGAATTTGGAGAGCTCTCCAAATTTATATTC 4860
Db |
4936 TTGAAAACCTCCCTAACCTATGCTCCAGGAATTTGGAGAGCTCTCCAAATTTATATTC 4995
Qy |
4861 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGCTTTGGAC 4920
Db |
4996 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGCTTTGGAC 5055
Qy |
4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCGCCAGCCGACATCTCA 4980
Db |
5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCGCCAGCCGACATCTCA 5115
Qy |
4981 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5040

Db |
5116 CCACAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5175
Qy |
5041 GGGCTCTGTCATCTGTGTCAGCATGCCAGCAAAAGTGTGCTATTCTCGCTACAAAGAAC 5100
Db |
5176 GGGCTCTGTCATCTGTGTCAGCATGCCAGCAAAAGTGTGCTATTCTCGCTACAAAGAAC 5235
Qy |
5101 CTAGGAAATATCTGATCCGGAAGAGATAGACCTCGAGCCCTGCGAGCTGATCCAC 5160
Db |
5236 CTAGGAAATATCTGATCCGGAAGAGATAGACCTCGAGCCCTGCGAGCTGATCCAC 5295
Qy |
5161 TTCAACAAATACAGTATCTCTCATTTGGAAACCAATAAATTTACGAAATCGACATGAAGCAG 5220
Db |
5296 TTCAACAAATACAGTATCTCTCATTTGGAAACCAATAAATTTACGAAATCGACATGAAGCAG 5355
Qy |
5221 TACAGCTCGAGGAATTCCTGATTAAGATGACATTCCTTGGCACCTGCTGTGTTGCC 5280
Db |
5356 TACAGCTCGAGGAATTCCTGATTAAGATGACATTCCTTGGCACCTGCTGTGTTGCC 5415
Qy |
5281 GCCTTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGAGGCGAGCAGAG 5340
Db |
5416 GCCTTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGAGGCGAGCAGAG 5475
Qy |
5341 GAGTACTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5400
Db |
5476 GAGTACTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535
Qy |
5401 CGCACAGCATCTCAAGTGGAGTTCCTTACCTTTGCGCTTTCCTTACAGAGAACCTAT 5460
Db |
5536 CGCACAGCATCTCAAGTGGAGTTCCTTACCTTTGCGCTTTCCTTACAGAGAACCTAT 5595
Qy |
5461 CTGTTTGTGAGCCCACTTCAACTCAGTGAATTTGAGATCCAGCAAGCTCCTCAGCA 5520
Db |
5596 CTGTTTGTGAGCCCACTTCAACTCAGTGAATTTGAGATCCAGCAAGCTCCTCAGCA 5655
Qy |
5521 GGGACCTCTCCGAGGCTACCTGGACATCCCAACCCGCTTACCTGGCCCTGCGCAT 5580
Db |
5656 GGGACCTCTCCGAGGCTACCTGGACATCCCAACCCGCTTACCTGGCCCTGCGCAT 5715
Qy |
5581 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAAATAAGGCTCATTTGCTGC 5640
Db |
5716 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAAATAAGGCTCATTTGCTGC 5775
Qy |
5641 AAGGAAACCTCTGTAAGAGTCCGACATCAACACCGGGCCGCTCCACCTCCGC 5700
Db |
5776 AAGGAAACCTCTGTAAGAGTCCGACATCAACACCGGGCCGCTCCACCTCCGC 5835
Qy |
5701 AG 5702
Db |
5836 AG 5837

RESULT 13
AF086824

LOCUS AF086824 6954 bp mRNA linear ROD 11-NOV-1998
DEFINITION Mus musculus rho/rac-interacting citron kinase (Crik) mRNA,

complete cds.

ACCESSION AF086824
VERSION AF086824.1 GI:3599508

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 6954)
AUTHORS Di Cunto,F., Calautti,E., Hsiao,J., Ong,L., Topley,G., Turco,E. and Dotto,G.P.

TITLE Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the Rho-Rac-binding protein Citron

J. Biol. Chem. 273 (45), 29706-29711 (1998)

JOURNAL 99009084
MEDLINE 9792683
PUBMED
REFERENCE 2 (bases 1 to 6954)

AUTHORS Di Cunto,F., Calautti,E., Hsiao,J., Ong,L., Topley,G., Turco,E. and Dotto,G.P.
 TITLE Direct Submission
 JOURNAL Submitted (25-AUG-1998) Biology, University of Torino, Via Santena 5 bis, Torino, TO 10126, Italy
 FEATURES Location/Qualifiers
 source
 1. .6954
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="spleen"
 1. .6954
 /gene="Crik"
 772. .6939
 /gene="Crik"
 /function="protein kinase"
 /note="CRK; rho/rac binding protein; alternatively spliced form of citron encoded by GenBank Accession Number AF039218; similar to Mus musculus citron-K encoded by GenBank Accession Number AF070066; similar to Rattus norvegicus citron-K encoded by GenBank Accession Number AF070065"
 /codon_start=1
 /product="rho/rac-interacting citron kinase"
 /protein_id="AAC72823.1"
 /db_xref="GI:3599509"
 /translation="MLKFKYGVNPPPEASBPFIASRASRLNLFQGGKPLMTQOOMS
 ALSREGLDALPALPECSOPALMKMKHVSFVKYSOTIAELRELQASDFEVRSL
 VQGFHFAEVVVEKATGDVYAKMKKALLAQVAFPEERNILSRSTPMTPOL
 YQAFQDKNLLVWYQPGDGLSILNRYEDQIDSMIQFYLAEILLAVSHVHGTV
 HDIPEKILLIDRTKIDLFQDSAAKNSNKKVDKPLIGTDPVAPVEVHEDRR
 GYGLDCDWVGVVAVMYGKTFTEGTSARTFNIMNFORFLKFPDDPKVSESL
 DILQSLCKQERLAFEGELCCHFPARTDMNNIRNPPFPVFLKSDSDTNSFDEPK
 NWSAFILCPSPKPLAPSGELPFVGFYSKALGYLGRSESVSSIDSPAKVSSMEKL
 LKSKELQSDPKCHKVEOMETRLHRRVSEAVLSQKEVELKASSETORSILLELAT
 VITECSLARSLEQRMVYSQDDKALQLLHDIHQSRKLQEIKEQVQAOVEEMLM
 MNQBEDLSARRSDLIPESELRESRLAEEFKRANSCQHKLMKAKQGGPEVGEYS
 KLEKINAEQLKIQELQEKARVASTATEALLQNIQAKERAREERLEKLNREDSS
 EGTKKLEAEERHLENKVRLEMERENRDLDDIQTSQEQVQADKILELEEK
 HREAQVSAHELVHKQREHYEETIKVLNDQIKDLADKESLENMQRHEEAHEKG
 KILSKQAMINAKDSIRLEOKIVELSEANKLAANSLSFTORNKAQEMISELRQK
 KYLEFQACKLAQNRKLEOLEKISHQSDSKRILLELETRLEVSLEHBEQKLELK
 ROLTELQSLQRESOLTAQARALESQLOAKTELEETAEAEESIQALTARDE
 IORFDALNRSTCTIDLEQNLTEDNLENNQNFYLSKOLDASANGEIVOLRS
 EVDHUREITREMOTLSQKQTEALKTCTLWLEQVLDLEALNDELLEKQEWAWR
 SVLGDSQFECRQELQMLDTEKQSRADQRTESQVVELAVKHEKABILALQ
 ALKEQKLAESDLNLEKXHAMLENARSLOKLETERELKORLEEEQAKLQOM
 DLQKNIHFRITQGLQALDRADLLKTERSDLEVQLENIQVLSHEKVMKEGTISOQTK
 LIDFLQAKDQPAKKKVPLOYNELKLALEKAKCAELEALQKTRIELRSAREEA
 HRKATHPHSPATARQIAMSIVRSEHQPASMSLLAPSSRKSSSTPEEFSRR
 LKERMHNIPHRNGLNRATKCAVCLDTVHFGRQASKLEQVNCBPKQCTCLPAT
 CGLPBYATHTEAFCDKNSGPGSQSEPSLSHLEGMKVFNRNKRQGGQWDERKI
 VLEGSKVLYDNEAREAGQRPVEEFELCLPDGDSVTHGAVASELANTAKADVPYILK
 MESHPHTTCWPGRTLLYLLAPSPDKQRTVALESVAVGRVREKAEADAKLLGNSLL
 KLEGGDRLDWNCITLPSDOVLVGTBEGLYALNVLNLSUTHIPGIVAVFOIYIKOLE
 KLMIAGEREALCLVDVKVQKSLAOSHLPADVPSPNIFEAVKSHLFAAGKIENSL
 CICAAMPKSVILRYNDNLKVCIRKEIETSPBCSIHFTNYSILGINKFYEIDMKQ
 YTLDEFLKNDHSLAPAVASSNSFPFVIVQANSAGQREYLLCFHFPFVDSYGR
 RSTDDLKWSRLFLAYREPFLVTFHNSLEVEIQARSSLSGPARYLEIPNRYL
 GPAISSGAIYLASQYDKLRVICCKGNLVKSETEQHRVPSRSRPNKRGPPTNEH
 ITRKRVASPPAPGSPHSPREPSPHRYDRREGRTELRDKSPGRPLERKSPGRMLST
 RRRSPGRFLPDESSRGLPAGAVITLPSQNVKWDQSSV"

ORIGIN

Query Match 70.5%; Score 4632.4; DB 10; Length 6954;
 Best Local Similarity 85.6%; Pred. No. 0;
 Matches 5341; Conservative 0; Mismatches 641; Indels 254; Gaps 7;
 QY 4 GCCGCCAGTGGGAGATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGATGCTGCT 63
 Db 757 GGGTTGGCGGAGAGATGTTGAAGTTCAGATATGTTGGCGGAACCGCGCGGAGGCAGT 816
 QY 64 GCTGCTGAACCCATTCGCCAGCGGGCTCCAGGCTGAATGTTCTTCAGGGGAAACCA 123

Db 817 GCCTCCGAGCCCATTCGCGGCTCCAGGCTAAATCTCTCTCCAGGGGAAACCG 876
 QY 124 CCCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAGAGGATATTAGATGCCTC 183
 Db 877 CCCCTCATGACTCAACAGCAGATGTCCTCTTTCCGGAAGGATGCTAGAGCCCTC 936
 QY 184 TTTCTTCTCTTTGAAGAATGCAGTCAGCTCTCTCTGATGAAGATTAAAGCAGCTGAGCAAC 243
 Db 937 TTGCTCTCTTTGAAGAGTGCAGCCACCCCTCTGATGAAGATGAAGCAGCTGAGCAGC 996
 QY 244 TTTGTCGGGAAGTATTCGGAACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAG 303
 Db 997 TTTGTCGAGAAGTATTCGGAACCATAGCTGAGTTGCGGAGCTGCGAGCCGTGGCGAGA 1056
 QY 304 GACTTCGAAGTCAGAAGTCTTGTAGTGTGTGTCACCTTCTGAGTGCAGGTGCTGAAGA 363
 Db 1057 GACTTCGAAGTTCGAAGCCCTTGTGGCTGTGTCTGCTCTGAGTGCAGGTGCTGAGA 1116
 QY 364 GAGAAAGCAACCGGGGACATCTATGCTATGAAAGTGTAGGAAGAAAGGCTTTATTGGCC 423
 Db 1117 GAGAAAGCGACCGGGGACGCTATGCTCCATGAAATCATGAAGAAGAAGGCTTTGCTGGCC 1176
 QY 424 CAGGAGCAGGTTTCATTTTGTAGGAGAGCGGGAACATATTATCTCGAAGCACAAGCCCG 483
 Db 1177 CAGGAACAGGTTTCATTTTTCGAGGAGGAGAGAACATATTATCTCGGAGCACGAGTCT 1236
 QY 484 TGGATCCCCCAATACAGTATGCTCTTTCAGACAAAAATCACCTTTATCTGATGAGGAA 543
 Db 1237 TGGATCCCCCAGTTACAGTACGCTTTTCAGGACAAAAATAACCTTTACCTGGTCATGAA 1296
 QY 544 TATCAGCTTGAGGGGACTGCTGTCTACTTTTGAATAGATATGAGGACCAAGTTAGTGAA 603
 Db 1297 TATCAGCTTGAGGGGATTTCTGTCTCTGTGAACAGATACGAGGACCAATTAGATGAG 1356
 QY 604 AACCTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTTCATCTGATG 663
 Db 1357 AGCATGATCAGTTTACCTTGTGCTGAGCTGATTTTGGCTGTCCACAGGTCACCATG 1416
 QY 664 GGATACGTGATCAGACATCAAGCCCTGAGAACATTTCTGTTGACCGCAGAGCACATC 723
 Db 1417 GGATATGTGATCAGACATCAAGCCCGAGAACATCTCATCGACCGGAGCGGAGATC 1476
 QY 724 AAGCTGTGATTTTGGATCTGCCCGGAAATGAATTCAAACAGATGTTGAATGCCAAA 783
 Db 1477 AAGCTGTGATTTTGGATCAGCCGCTAGANTGAATTCAAATAA ---GGTGGATGCCAAA 1533
 QY 784 CTCCCGATTGGGACCCCGAGATTACATGGCTCCTGAACTGCTGATGTGATGAACGGGAT 843
 Db 1534 CTCCCATTTGGGACCCCGGATTACATGGCTCCGGAAGTGTGACCGGTGATGAACGAGAC 1593
 QY 844 GGAAGGACACTACGGCTGGACTGTGACTGGTGGTCACTGGGCGTGAATGCCTATGAG 903
 Db 1594 CGAAGGGGACATACGGCTTGGACTGTGACTGGTGGTCTGTGCGGAGTGTGCTTATGAG 1653
 QY 904 ATGATTTATGGGAGATCCCTCTCGCAGAGGGAACCTCTGCGAGAACCTTCAATAACATT 963
 Db 1654 ATGTTTATGGGAAGACCCCATTCACAGAGGGAACCTTCGCGCGGACCTTCAACAACATC 1713
 QY 964 ATGAATTTCCAGCGGTTTGAATAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAATTTCT 1023
 Db 1714 ATGAATTTCCAGCGGTTTGAAGTTCACATGATGACCCCAAGTGAAGTGAAGTGAATTTCT 1773
 QY 1024 GATCTGATTCAAAGCTTGTGTGCGCGCAGAAAGAGAGACTGAAGTTTGAAGTGTCTTTCG 1083
 Db 1774 GATCTGTTTCAAGTGTGTGTGTGTCAGAAAGAGAGACTGAAGTTTGAAGTGTCTTTCG 1833
 QY 1084 TGCCATCCTTTCTCTCTAAATTTGACTGGAAACAATTCGTAATCTCTCTCCCTCCCTTC 1143
 Db 1834 TGCCACCCCTTTCTTTTGGCAGAACCGACTGGAAACAACATCCGTAATCTCTCTCCCTCCCTTC 1893
 QY 1144 GTTCCACCCCTCAAGTCTGACGATGACACCTCCCAATTTTGAATGATGACAGAGAAATTCG 1203

3315	Qy	GAGCAGGGGAGAGCGGATCAGCGGATACCGGAGTCTCGCCAGGTGGTGGAGCTGGCAGT	3374
4113	Db	GAGCAGGGCTAGGGCCGATCAGCGGATACCGGATCGCGCAGGTGGTGGAGTTGGCGGT	4172
3375	Qy	GAAGGAGACAAGGCTGAGATTCTCGCTCTGCAGCAGGCTCTCAAAGACAGAAAGCTGAA	3434
4173	Db	GAAGGACACAGGCGGAGATTCTTGCTCTGCAGCAGGCTCTCAAGGACGAGACTCAA	4232
3435	Qy	GGCCGAGAGCCTCTTGACAAGCTCAATGACTTGGAGAGAAGCATGCTATGCTTGAAT	3494
4233	Db	AGCCGAGAGCCTGTGGACAAAGCTCAACGACCTTGGAGAGAAGACGCCATGTCTGGAGAT	4292
3495	Qy	GAATGCCCGAAGCTTACAGCAGAAAGCTGGAGACTGAACGAGAGCTCAACAGAGGTTCT	3554
4293	Db	GAACGCCCGAGCTTACAGCAGAAATCTAGAGACAGAGCGGAGCTCAACAGAGGTTCT	4352
3555	Qy	GGAGAGCAAGCCAAATTACAGCAGCAGATGGACCTCGAGAAAAATCACATTTCCGTCT	3614
4353	Db	GGAGGAGCAAGCCAAATTGAGCAGCAGATGGACCTGCAGAGAAACCACTCTTCAGACT	4412
3615	Qy	GACTCAAGGACTCGAAGAAGCTCTAGATCGGCTGATCTACTTGAAGACAGAAAGAGTGA	3674
4413	Db	GAGCGAAGGCTGACGAGGCGCTGGACCCGGGCGGATCTGCTGAAGACAGAAAGGAGCGA	4472
3675	Qy	CTTGGAGTATCAGCTGGAAAAACATTACGGTTCTCTATTCTCATGAAAGGTGAAATGGA	3734
4473	Db	CTTGGAGTACGACTGGAACAATTCAGGTTCTCTACTCTACAGAGAAAGTGAATGGA	4532
3735	Qy	AGGCACTATTCTCAACAAACCAAATCTCATTTGATTTTCTCAAGCCAAATGACCAACC	3794
4533	Db	AGGCACAATCTCTCAGCAAAACCAACTCATTTGATTTCTCTCAAGCCAAATGACCAACC	4592
3795	Qy	TGCTAAAAAGAAAAGGGTTTATTATTAGTCAGCGAAAGAGGACCTTGCTTTACCCACACA	3854
4593	Db	TGCTAAAAAGAAA-----A	4607
3855	Qy	GGTTCTCTCGAGTACAATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGC	3914
4608	Db	GGTTCTCTCGAGTACAATGAGCTGAAGCTAGCCCTGGAAAAGAGAAAAGCCCGATGCGC	4667
3915	Qy	AGAGCTAGAGGAAGCCCTTCAGAGAGCCGCATCGAGCTCCGGTCTGCCCGGAGGAAGC	3974
4668	Db	GGAGCTGGAGGAGGCCCTTCAGAGAGCCGCATCGAACTCGCTCTGCGCGGAGGAAGC	4727
3975	Qy	TGCCACCGCAAAAGCAACGGACCAACCAACCCATCCACGCCAGCCACCGCAGGAGCA	4034
4728	Db	TGCCACCGCAAAAGCCACGGACCAACCGCACCCGCTCTACGCCACCTCGCAGGAGCA	4787
4035	Qy	GATCGCCATGTCCGCATCTGCGGTCCAGAGACACAGCCAGCCAGTGCATGAGCCCTGCT	4094
4788	Db	GATCGCCATGTCTGCCATTGTGGGTTCGCCGAGCACAGCCAGTGCCTATGAGCCTGCT	4847
4095	Qy	GGCCCCGCCATCCAGCCGCGAAGAGGAGTCTTAACTCCAGAGAAATTTAGTGGCGCTCT	4154
4848	Db	TGCCCCACCATCCAGCCGCGAAGAGGATGCTCAACTCTAGAGAAATTCAGCCGCGCTCT	4907
4155	Qy	TAAGGAACGATGCAACCACAATATTCCTCACCGATTCAAGTAGGACTGAACATGCGAGC	4214
4908	Db	GAAGAGCGCATGACCAACAATCTCCCTACCCGTTTAAATGTGGGCGCTGAACATGAGAGC	4967
4215	Qy	CACAAAGTGTGCTGTGTCTGGATACGGTSCACTTTGGAGCGCCAGAGCATCCAAATGCTCT	4274
4968	Db	CACCAAGTGCGCCGTGTGTCTGGATATCTGTGCATTTTGGAGCGCCAGCGGATCCAAATGCTCT	5027
4275	Qy	CGAATGTCAAGTGTGTGTCAACCCAAAGTGTCCACGTGCTTGCCAGCCACCTGCGGCTT	4334
5028	Db	CGAATGTCAAGTGTGTGTCAACCCAAATGTCTCACCTGCTTGGCCGCTACCTGTGCGCT	5087
4335	Qy	GCCTGCTGAATATGCCACAGACTTCCCGAGGCGCTTCTGCGGTGACAAAATGACTCTCCC	4394
5088	Db	GCCAGCTGAATATGCCACAGACTTCACTGAGCGCTTCTGCGGGACAAAATGAACTCCCC	5147
4395	Qy	AGGTCTCCAGAACCAAGAGCGCAGCAGAGCTTGCACTGTGAAGGGGTGGATGAAGTGCC	4454

Db	5148	GGGGCTCCAGAGCAAGAGCCTGGCAGCAGCTTGACCTGAGAGGTGGATGAAGTGCC	5207
Qy	4455	CAGGAATAACAACAGCAGCAGCAGCAAGCTCGGGAAGAGTACATTTCTCTTGAGGGATC	4514
Db	5208	CAGAAATAACAACCGGGAAGCAGCAAGCTCGGGAAGAGTACATTTCTCTTGAGGGGTC	5267
Qy	4515	AAAAGTCCTCATTTATCACAATGAAGCCAGAGAAGCTGGACAGAGGCCGCTGGAAGAATT	4574
Db	5268	AAAAGTCCTCATTTATCACAATGAAGCCAGAGAAGCTGGACAGAGGCCGCTGGAAGAATT	5327
Qy	4575	TGAGCTGTGCCTTCCGACGGGATGTATCTATTATGTGTGCCGTGGTCTCCGAACT	4634
Db	5328	TGAGCTGTGCCTTCCGACGGGACGTATCTATTATGTGTGCCGTGGTCTCCGAACT	5387
Qy	4635	CGCAAAATACAGCCAAAGCA-----	4653
Db	5388	TGCAAAATACCGCCAAAGCAGATGTCCATATCTCTGAAAGATGAGTCTCACCCACAC	5447
Qy	4654	-----	4653
Db	5448	CACCTGTGCGCTGGGAAACCTCTACTTGTGTAGCACCCAGCTTCCCGACAAGCAGCG	5507
Qy	4654	-----GAAAAAGCAGA	4664
Db	5508	CTGGGTCAACGCGCTTAGAATCTGTCTGCGAGGTGGAGAGTTTCTAGGGAAAAAGCCGA	5567
Qy	4665	AGCTGATGCTAAATCTGCTTGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGA	4724
Db	5568	AGCCGATGCTAAATCTCTTGAAACTCTCTGCTGAAACTGAGAGGCCGATGACCGGCTTGA	5627
Qy	4725	CATGAACTGCAAGCTGCGCTTCACTGACCAAGTGGTGTGGTGGGCACCCGAGGAAGGCT	4784
Db	5628	CATGAACTGCAAGCTGCGCTTCACTGACCAAGTGGTGTGGTGGGCACCCGAGGAAGGCT	5687
Qy	4785	CTAGCCTCTGAATCTCTTGAAAACTCCCTAAACCCATGCTCCAGGAATTTGGAGCAGTCT	4844
Db	5688	CTAGCCTCTGAATCTCTTGAAAACTCCCTAAACCCATGCTCCAGGAATTTGGCGCAGTCT	5747
Qy	4845	CCAAATTTATATTCAAGGACCTGGAGAAGCTACTCATGATAGCAGAGAAGCGCGC	4904
Db	5748	CCAAATTTATCATCAAGGACCTGGAGAAGCTCTCATGATAGCAGGGAAGAGCGCGC	5807
Qy	4905	ACTGTGCTTTGAGCATGAAGAAGTGAAACAGTCCCTGGCCAGTCCCACTCGCTGCCTGC	4964
Db	5808	TCTGTGCCCTGGTGGACGTGAAGAAGTGAAACAGTCCCTGGCCAGTCCCACTCGCTGC	5867
Qy	4965	CCAGCCGACATCTCACCCAACTTTTGAAGCTGTCAAAGGCTGCCACTTTGTTGGGGC	5024
Db	5868	CCAGCCGACGCTCTCCCAACATATTCGAAGCCGTCAAAGGCTGCCACTTTGTTGGTGC	5927
Qy	5025	AGGCAAGATTGAGAAACGGGCTCTGCATCTGTGCAAGCATGCCCAAGCAAAAGTCGTCTATCT	5084
Db	5928	TGGCAAGATCAGAACACGCTTGTCATCTGCGCGCTATGCAAGCAAGAGTCGTCTATCCT	5987
Qy	5085	CCGCTACAAACAAACCTCAGCAAAATCTGCATCCGGAAGAGAGATAGAGCCTCAGAGCC	5144
Db	5988	CCGCTACAAAGCAACCTCAGCAAGTACTGCATCCGCAAGGAGATCGAGACCTCAGAGCC	6047
Qy	5145	CTGCAGCTGTATCCACTTCAACAAATTACAGTATCCTCATTTGGAAACCAATAAATTCTAGA	5204
Db	6048	CTGCAGCTGTATCCACTTCAACAAATTACAGTATCCTCATTTGGAAACCAATAAATTCTAGA	6107
Qy	5205	AATCGACATGAAGCAGTACAGCTCGAGGAATTCCTGGTAAGAATGACCACTTCTTGGC	5264
Db	6108	GATCGACATGAAGCAGTACAGCTTGAAGTTCCTGGCAAGAAACCAATTCCTTGGC	6167
Qy	5265	ACCTGCTGTGTTGCGCGCTCTTCCAAACAGCTTCCCTGTCTCAATCTGTCAGAGTGAACAG	5324
Db	6168	ACCTGCTGTGTTGCGCTCTTCCAAACAGCTTCCCTGTCTCAATCTGTCAGAGTGAACAG	6227
Qy	5325	CGCAGGCGAGCAGGAGTACTTGTGTGTTTCCAGAAATTTGGAGTGTCTGTTGATTC	5384

6228 GCCGGGAGGAGAGAAATACCTGCTGTGCTTCCAGAAATTTGGGGTGTTCGTGGATTC 6287
QY 5385 TTACGGAAGACGTAGCCGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCTTTGC 5444
Db 6288 TTACGGAAGACGTAGCCGACAGATGATCTCAAGTGGAGTGCCTTACCTCTGGCTTCGC 6347
QY 5445 CTACAGAGAACCTATCTGTTTGTGACCCACTTCAACTCACTCAAGTAATAGATCCA 5504
Db 6348 CTACAGAGAACCTATCTGTTTGTGACCTCACTCAACTCCCTGGAAGTCAATGAGATCCA 6407
QY 5505 GGCACGCTCTCCAGCAGGAGCCCTGCCGAGCGCTACTCGACATCCGCAACCGCGCTA 5564
Db 6408 GGCAGATCTCTCACTGGGAGCCCTGCCGAGCATATCTGGAATTTCAAACTCTGCTA 6467
QY 5565 CTGGGCGCTCCGATTTCTCTCAGAGCGATTTACTTGGCGTCTCTATACAGGATAAAT 5624
Db 6468 CTGGGCGCTCCGATTTCTCTCAGAGCGATTTACTTGGCGTCTCTATACAGGATCAAT 6527
QY 5625 AAGGCTCATTTGCTGCAAGGAGAACCTCTGAGAGGCTCCGGCACTGAACACACCGGG 5684
Db 6528 AAGGCTCATTTGCTGCAAGGAGAACCTCTGAGAGGCTCCGGCACTGAAGGCTGAGCAGCCGGT 6587
QY 5685 CCGTCCACCTCCGAGAGCCGACCAAGAGGAGCCGACCCAGCTGACCAAGAGCAGCAT 5744
Db 6588 GCGCTCCACCTCCGAGAGCCGACCAAGAGGAGCCGACCCAGCTGACCAAGAGCAGCAT 6647
QY 5745 CACCAAGCGGTGCTCCAGCCAGCGCCGCGGAGGCGCCGAGGCGCCAGCCGCGAGGCC 5804
Db 6648 CACCAAGCGGTGCTCCAGCGCGGCGGAGGCGCCGAGGCGCCAGGCGCCGAGGCGCC 6707
QY 5805 AAGCACACCCACCGCTACCC-----CGCAGGCGGCGAGCGCTGCGCAGGAGCAAGTC 5858
Db 6708 AAGCACACCCACCGCTACCCAGAGACAGAGAGGCTCGGACAGAGCTCGGAGGAGCAAGTC 6767
QY 5859 TCCTGGCGCGCTCTGAGGAGAGAGAGTCTCCCGCGCGGATGCTCAGACGCGGAGAGA 5918
Db 6768 TCAGGCGCGCTCTGAGGAGGAGAGTCTCCCGCGCGGATGCTCAGACACTAGGAGAGA 6827
QY 5919 GCGTCTCCCGGAGGCTGTTTGAAGACAGCAGCGGCGCGCTGCGGAGGCGCT 5978
Db 6828 GCGTCTCCCGGAGGCTGTTTGAAGACAGCAGCGGCGCGCTGCGGAGGCGCT 5887
QY 5979 GAGACACCCGCTGCTCCAGGTGAACAGGAGGAGGCGAGGCTCTCTCAAGTTT 6034
Db 6888 GAGACACCCGCTGCTCCAGGTGAACAGGAGGCTGCGGAGGAGGCTCTCTCAAGTTT 6943

RESULT 14

AF039218 LOCUS AF039218 5952 bp mRNA linear ROD 02-JUL-1998
DEFINITION Rattus norvegicus postsynaptic density protein (citron) mRNA,
complete cds.
ACCESSION AF039218
VERSION AF039218.1
KEYWORDS GI:2745839

SOURCE

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 5952)
Zhang, W., Apperson, M.L., Vasquez, L.E. and Kennedy, M.B.
Citron, a PSD-95-binding protein at glutamatergic synapses on
inhibitory neurons
Unpublished
2 (bases 1 to 5952)
Zhang, W., Apperson, M.L. and Kennedy, M.B.
Direct Submission
Submitted (18-DEC-1997) Biology, Caltech, Pasadena, CA 91125, USA
JOURNAL Location/Qualifiers
FEATURES

source

1. .5952
/organism="Rattus norvegicus"
/mol_type="mRNA"

gene
CDS
/db_xref="taxon:10116"
1. .5952
/gene="citron"
613. .5469
/gene="citron"
/function="interacting with Rho family GTPases"
/codon_start=1
/product="postsynaptic density protein"
/protein_id="AAC25483.1"
/db_xref="GI:2745840"
/translation="MEQEMARLHRRVSEAVLSQEKVELKASQTSRLLEODLATYI
TECSLKSLSQARMEYSQEDDAXALQLLHDIREQSRKLOETKEQYQAOVEMRLMMN
KESDLVSARRRSDLYSELSRESRLAAEFKRAKANECHQKMKAKDLGKPEVGECSRL
QINAKOQLK:QELQKLEKAVKASTATELLQNIQAKERAERELEKJHNRDSSG
IKKLVAAEERRHSLNKVRELETMERENRDKDDIQTSQEQIQOMADKILEBEKHR
BAQVSAEHLVHLKQKQHYBEKIKVLDNQIKDLADKESLTMQSHBEAHEKGI
LSEQKAMNAMDKIRLSLEQRIVELSANKLAANSLSLTORNKMAQEMTSELQKQF
YLEQAGKLEAQNKLREOLEKISHQDSKRNLELETRLRVSLRSHERQKLELQ
LTELQSLQERESQITALQARALESQLRQAKTELBETTAEBEETQALTAHDEIQ
KRFADNRSTVITDLEEQNLQUTEDNAELNNQNFYLSKQIDEASGANDIVQLRSEV
DHLREITREMQLTSQKQMEALTKTTMLDEQVMDLEALNDELLEKQWEARSV
LGDEKQFECRVELQMLDTEKQSRARADQITESTQVVELAVKEHKAIBALQOQAL
KEQKKAESLDKLDLEKKHAMLENWARSLOQKLETERELKORLEBEQAKLOQMDL
QKNIIFRLTOGLBALDRADLLKTERSDLEVOLENIQVLYSHKVKVEGITSQTKLI
DFLOAKVDPAKKKGLFSRRKEDPALPTQVPLQYNELKLAKEKAKCAELRALQK
TRIELRSAREBAHRKATDHPHSTPATARQIQIAMSALVRSPEHOPMSALLAPPSSR
KREASTPEERSRLUKERHNIHPRFNVGLNMRATKCAVCLDITVHFGQKASKCLEQV
MCHPKSTCTPATCGLPAEYATHTEAFPCRDQKVSPPGLQSKESPSSUHLGKWKVPRN
NKGQQQWDRKVIYLEGSKVLIYDNEAREAGQPVVEFELCLPDGDSVSHGAVGASEL
ANTAKAPVYIKMESHPTTCMPGRTLLYLAQSPDKQWMTALSVWAGGRVSRK
AADAANDKSCBRLPVWVKLGNLSLLKLEGGDDRLDMNCTLPFSDQVVLVGTBEGLY
ALNVKLTHI PGIGAVFOIYIKLEKLLMIAGERALCLVDVKVKQSLAQSHLP
AQPVSNINFEAVKGCCHLPAAGKIENSLCICAAKPSKVILIRNDNLKSKIKETET
SEPCSHFNYSILIGNTKYQYEDMKQYLTLEFLDKNDHSLAPAVFASNTNFFVSI
VQANSTQREYELLCHEFGVFDVSGRRSTDDLKWSRLPLAFAYREPLYFTFNS
LEVEIQARSLGTPARAYLEINPRYLIPAISSGAIYLAASSYQDKLRLVCCCKNLVK
ESGTEHRVSTSRPNKRGPTVYNIHKRVASSPAPPEGSHPEPSPHVRDRE
GRTLEDRDKSPGRLEBREKSPGRMLSTRERSPGRLEDFEDSSRGLPAGAVRTPLSVQV
KWDQDSV"

misc_feature 3052..3096
/gene="citron"
/note="may be excised by alternative splicing"

misc_feature 4066..4110
/gene="citron"
/note="may be excised by alternative splicing"

ORIGIN

Query Match 54.8%; Score 3602.6; DB 10; Length 5952;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 4171; Conservative 0; Mismatches 444; Indels 252; Gaps 5;
QY 1417 AAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTACAGAGTGGAGGCTTGCTT 1476
Db 610 ATGATGAGCAGGAAATGCGCCGGTTCATCGCAGAGTGTACAGAGTGGAGGCTTGCTT 659
QY 1477 AGTCAGAGGAGTGGAGCTGAAGGCGCTCTGAGACTCAGAGATCCCTCTCGGAGCAGGAC 1536
Db 670 AGTCAGAGGAGTGGAGCTGAAGGCGCTCTGAGACTCAGAGATCCCTCTCGGAGCAGGAC 729
QY 1537 CTTCCTACCTACATCAGAGATCGAGTCTTAAGCGAGTTTGGAGCAAGCAGCATG 1596
Db 730 CTTCCTACCTACATCAGAGATCGAGTCTTAAGCGAGTTTGGAGCAAGCAGCATG 789
QY 1597 GAGTGTCTCCAGGAGGATGACAAAGCAGCTCTCCATCATATCAGAGACGAGC 1656
Db 790 GAGTGTCTCCAGGAGGATGACAAAGCAGCTCTCCATCATATCAGAGACGAGC 849
QY 1657 CGGAGCTCCAGAGATCAAGAGCAGGAGTACAGGCTCAAGTGAAGAAATCAGGTTG 1716
Db 850 CGGAGCTCCAGAGATCAAGAGCAGGAGTACAGGCTCAAGTGAAGAAATCAGGTTG 909
QY 1717 ATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGACGAGTGTATCTCTACGAA 1776

Db 910 ATGATGAATCAGCTGGAAGAGGACCTCTGTCTCAGCCCGAAGACGAGCGATCTCTACGAG 969
QY 1777 TCTGAGCTCAGAGAGTCTCGGTTGCTGCTGAAGAATTCAACGGGAACGACGAAATGT 1836
Db 970 TCTGAGCTCAGGAGTCTCGGTTGCTGCGCCGAGGAATTCAAGCGGAAGCAAAACGAAATGT 1029
QY 1837 CAGCATAAACTTTGAAGCTTAAGGATCAAGGGAAGCCTGAAGTGGGAAATATCGGAA 1896
Db 1030 CAGCATAAACTGATGAAGCTTAAGGACCTAGGGAAGCCTGAAGTGGGAAATATGTTCCAGA 1089
QY 1897 CTGGAGAAGATCAATGCTCAGCAGCAGCTCAAAATTCAGGAGCTCAAGAGAAACTGGAG 1956
Db 1090 CTGGAGAAGATCAATGCTCAGCAGCAGCTGAAGATTCAGGAGCTCCAGGAGAAGCTGGAG 1149
QY 1957 AAGGCT-----GCAAG 1968
Db 1150 AAGGCTGTAAAGCCAGCACCGAGGCCACAGAGCTCTCTCGAAGACATCCGGCAGGCAAG 1209
QY 1969 GAGCGAGCCGAGGAGGAGCTGGAGAAGCTGCAGAACCGAGAGGATTTCTTGAAGGCATC 2028
Db 1210 GAGCGAGCTGAGGAGGAGCTGGAGAAGCTGCACACAGGGAAGATTTCTCGAAGGCATC 1269
QY 2029 AGAAGAAGCTGTTGAAGCTGAGGAACCGCGCATTTCTCTGGAGAAACAAGGTAAGAGA 2088
Db 1270 AAAAGAAGACTGTTGAAGCTGAGGAACCGCGCATCTCCCTGGAGAAACAAGGTAAGAGA 1329
QY 2089 CTAGAGACCATGAGCGGTAGAGAAACACAGACTGAAGGATGACATCCAGACAAATCCCAA 2148
Db 1330 CTAGAGACCATGAGCGGTAGAGAAACACAGACTGAAGGATGACATCCAGACAAATCCGAA 1389
QY 2149 CAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAAGAGAAACATCGGAGGCCAA 2208
Db 1390 CAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTGGAGGAAGCATCGGAGGCTCAG 1449
QY 2209 GTCTAGCCCGACCTAGAGTGCACCTGAAACAGAGAGACGACACTATCAGAGAAAG 2268
Db 1450 GTCTAGCTCAACACTGAAGTACACCTGAAAGCAGAGGAACAGCACTATCAGAGAAAG 1509
QY 2269 ATTAAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAC 2328
Db 1510 ATCAAGTGTGGACAATCAGATAAAGAAAGACTTGGCCGACAGAGAGGCTTAGAGACC 1569
QY 2329 ATGATCCAGAGACAGAGGAGGCGCATGAGAGGGCAAAATTTCTAGCGAAACAGAG 2388
Db 1570 ATGATCAGAGGCACAGAGAGGCGCCACGAGAGGGCAAGATCCTCAGCGAGCAGAG 1629
QY 2389 GCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGGAAGAGGATTTGTGAGCTG 2448
Db 1630 GCGATGATCAACGCAATGATTCAGATCCAGATCAGATCCCTGGAAGAGGATTTGTGAGCTG 1689
QY 2449 TCTGAAGCCAAATAAATTCAGCAAAATAGCAGTCTTTTACCCAAAGGAAACATGAAGGCC 2508
Db 1690 TCAGAAGCCAAACAAGCTTGGGCAACAGCAGTCTCTTACCCAGAGGAACATGAAGGCC 1749
QY 2509 CAAGAGAGATGATTTCTGAATCAGGCAACAGAAATTTTACCTGGAGACAGCTGGG 2568
Db 1750 CAGGAAGAGATGATCTCAGAACTCAGGCGAGCAAAATTTTACCTGGAGACACAGGCCGGG 1809
QY 2569 AAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTCGAGAAGATCAGCCACCAAGAC 2628
Db 1810 AAGCTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTAGAGAGATCAGCCACCAAGAT 1869
QY 2629 CACAGTGACAGAAATCGGCTGTGGAATCTGGAGACAAAGATTGGGGAGGTGATCTAGAG 2688
Db 1870 CACAGTGACAAAGATCGGCTGTGGAGCTGGAGCAAAAGTTGAGAGGCTCAGCCTGGAG 1929
QY 2689 CACGAGGAGCAGAAATCGGAGCTCAAGCGCCAGCTCAGAGGCTCAGGCTCTCCCTGGCAG 2748
Db 1930 CATGAGAGCGAAGCTGAGACTGAAGCGGACGCTCAGCGAGCTGCAGGCTCTCCCTGCGAG 1989
QY 2749 GAGCGGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGCGCGCCCTGGAGAGCAGCTT 2808
Db 1990 GAGCGTGAAGTCCAGAGCTGAGCGGCTGAGGCGCGCCGCGGCTGCCCTTGAGAGCAGCTC 2049

QY 2809 CGCCAGCGGGAAGACAGAGCTGGAAAGAGACACACAGCAGAAAGCTGAAGAGGAGATCCAGGCA 2868
Db 2050 CGCCAGCGGGAAGACAGGAGCTGGAGGAGACAACTGCAGAAAGCGGAGGAGATCCAGGCA 2109
QY 2869 CTCACGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACT 2928
Db 2110 CTCACGGCACATCGAGATGAATCCAGCGCAAAATTCGATGCCCTTCGCAACAGCTGTACT 2169
QY 2929 GTAATCAGACACTGGAGGAGCAGCTAAACAGCTGACCGAGGACAAACGCTGAATCAAC 2988
Db 2170 GTCATCACCGACCTGGAGGAGCAGCTGAACAGCTCAACGAGGACAAACGCGGAGCTCAAC 2229
QY 2989 AACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGCGCCAAACAGCAGAT 3048
Db 2230 AACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGCGTGCACAGATGAT 2289
QY 3049 GTAACACTGGAAGTGAAGTGGACCATCTCGCGGGAGATCAACGGAACAGAGATCGAG 3108
Db 2290 GTGCAGCTGCGAAGCGAGGTGGACCATCTCGCGCTGAGATCACGAGCGGAGATCGAG 2349
QY 3109 CTTACCGCCAGAAAGCAACGATGGAGCTTGAAGACCACTGTCACCATCTCTGGAGAA 3168
Db 2350 CTCACCGCCAGAAAGCAACATGGAGCTTCTGAAGACGACCTGCACCATCTCTGGAGAG 2409
QY 3169 CAGGTCATGGATTTGGAGGCGCTAAACGATGAGCTGTAGAAAAGAGCGGCACTGGGAG 3228
Db 2410 CAGGTCATGGACTTGGAGGCGCTGAAACGACGAGCTGTGGAAGAGGAGCGCAGTGGAG 2469
QY 3229 GCCTGGAGGAGCTCTCGGTTGATGAGAAATCCAGTTTGAAGTCTCGGTTTCGAGAGCTG 3288
Db 2470 GCCTGGAGGAGCTCTCGGCGGACGAGAGTCTCAGTTTGAAGTCTCGAGTTTCGAGAGCTA 2529
QY 3289 CAGAGATGCTGGACACCGAGAAACAGAGAGGCGGAGAGCCGATCAGCGGATCAGCGAG 3348
Db 2530 CAGAGATGCTGGACACCGAGAGACAGAGAGGCTTAGAGCTGATCAGCGGATCAGCGAG 2589
QY 3349 TCTCGCCAGTGTGGAGCTGGCAGTGAAGAGACACAAGGCTGAGATTTCTGCTCTCGAG 3408
Db 2590 TCTCGCCAGTGTGGAGCTGGCAGTGAAGAGACACAAGCGAGATTTCTCGCGCTCGAG 2649
QY 3409 CAGGCTCTCAAGAGCAGAGAGCTGAAGCGCGAGAGCCTCTCTGACAGCTCAATGAGCTG 3468
Db 2650 CAGGCTCTCAAGAGCAGAGAGCTGAAGAGCTGAGAGCTCTCGGACAAAGCTCAAGAGCTG 2709
QY 3469 GAGAGAGCAGTGTATGCTTGAATGAATGCCCGAAGCTTACAGCAGAAAGCTGGAGACT 3528
Db 2710 GAGAGAGCAGCGCATGCTTGAGATGAATGCCCGGAGCTTACAGCAGAACTGGAGACA 2769
QY 3529 GAAACGAGGCTCAACAGAGGCTTCTGGAAGAGCAGCGCAAAATTTACAGCAGCAGATGGAC 3588
Db 2770 GAGCGGAGCTCAACAGAGGCTTCTGGAGGAGCAAGCCAAGTTGCGAGCAGCAGATGGAC 2829
QY 3589 CTGCAGAAAAATCACATTTTCCGTCTGACTCAAGGACTGCAAGAGCTCTAGATCGGGCT 3648
Db 2830 CTGCAGAAAGAACCACTCTCAGACTGACACAAGGCTGCGAGAGGCGCTGGACCGGGCG 2889
QY 3649 GATCTAATGAAGAAGAGAGTGAATTTGAGATGATCAGCTGGAAGAAACATTCAGGTTCTC 3708
Db 2890 GATCTGCTGAAGACCGAAGAGGAGCGCTGGAATATCAGCTGGAAGAAACATTCAGGTTCTC 2949
QY 3709 TATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGAT 3768
Db 2950 TACTCACAGAGAGGTGAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGAT 3009
QY 3769 TTTCTGCAAGCCAAATGGACCAACCTGCTTAAAGAAAAAGGGTTTTATTTAGTCGACGG 3828
Db 3010 TTTCTGCAAGCCAAATGGACCAACCTGCTTAAAGAAAAAGGGTTTTATTTAGTCGACGG 3069
QY 3829 AAAGAGAGCCTGCTTTACCCACAGAGTTCTCTGCAAGTACATGAGCTGAAGCTGGCC 3888
Db 3070 AAAGAGAGCCTGCTTTGCCACACAGGTTCTCTGCAAGTACATGAGCTGAAGCTGGCC 3129

Db 5287 AGGACAGAGTCTCCAGGCGTCTCTGAGCGGAGAGTCCGACGCGGATGCTCAGC 5346

Qy 5908 ACCGAGAGAGCGTCCCGGAGGCTGTTTGAACAGACAGCAGCAGGCGCGGTGCT 5967

Db 5347 ACAGAGAGAGCGGTCGCGGAGACGTGTTGAAGACAGCAGCAGGCGCGGTGCT 5406

Qy 5968 GCGGAGCGGTGAGGACCGCGTGTCCAGGTTGAACAGGAGAGGCGCAGAGTGCCTCT 6027

Db 5407 GCGGAGCGGTGAGGACCGCACTGTCCAGGTTAACAAGTCTGGGACGAGTCTCAGTA 5466

Qy 6028 CAGTTT 6034

Db 5467 TAGTCT 5473

RESULT 15

AF070066

LOCUS Mus musculus Citron-K kinase mRNA, partial cds. ROD 31-JUL-1998

DEFINITION Mus musculus Citron-K kinase mRNA, partial cds.

ACCESSION AF070066

VERSION AF070066.1 GI:3360513

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 4967)

AUTHORS Madaule, P., Fuyushiki, T., Reid, T., Ishizaki, T., Watanabe, G., Morii, N. and Narumiya, S.

TITLE A novel partner for the GTP-bound forms of rho and rac

JOURNAL FEBS Lett. 377 (2), 243-248 (1995)

MEDLINE 96128238

PUBMED 8543060

REFERENCE 2 (bases 1 to 4967)

AUTHORS Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuo, T., Bito, H., Ishizaki, T. and Narumiya, S.

TITLE Role of Citron kinase as a target of the small GTPase Rho in cytokinesis

JOURNAL Nature (1998) In press

REFERENCE 3 (bases 1 to 4967)

AUTHORS Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuo, T., Bito, H., Ishizaki, T. and Narumiya, S.

TITLE Direct Submission

JOURNAL Submitted (04-JUN-1998) Pharmacology, Faculty of Medicine, Kyoto University, Yoshida-Konohe, Sakyo-ku, Kyoto 606 8315, Japan

FEATURES

source

1..4967

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="adult"

<1..4928

/function="binds to Rho and regulates cytokinesis"

/notes="kinase variant form of Citron; approximately the first 50 amino acids belong to the kinase domain; alternatively spliced form of Citron, encoded by GenBank Accession Number AF039218; similar to Rattus norvegicus Citron-K encoded by GenBank Accession Number AF070065"

/codon_start=3

/product="Citron-K kinase"

/protein_id="AAC27933.1"

/db_xref="GI:3360514"

/translation="PEVPTLKSDDTNFDPEKSNVSVVCSVQLSPGSGFELPFV

GRSVKALGYLGRSVSVSLDSPPAKVSMKELLIKSKELQSDQKCHKMEQVTRL

HRVSVAVLVSKVELKASFTQRLSLQDLATYITCSLSKRLSEQARMEVSED

KALQLLIREQSKLOETKEQYQAVEMRMNQLNEEDLVARSRLDYSELRE

SRLAEPFRKNECHQKLMKADQGPVGYGSKLEKINAEQQLKIQELQKLEKAV

KASTEALLQNIROAKERAERELEKLNHNSDSGEEKKLVEASELEKHEAQA

QHLVELKQEHVEKI KVLNNOIKKDLADKESLENMQRHEEAEHKGKILSQKA

MINAMSKTSRLSEQRVELSEBANKLAANSLEFQNMKAQEMISELRQKQYELTQA

KLEAFNRKLEQLEKIRSHQSDSKRLLELETRVSELEHBEKLEKRLQTLLEQL

SLQERESQUTALQAPAALESQLRQAKTELETTABEEI QALTARHDEIQKQFDAL

RNSCTVITDLEBOLNLTEDNAELNNFYLKQLEASGANDEI VOLRSEVDHRLRE

ITREMOITLTKOTMEALKTTCTMLEQVLDLEALNDELLEKQEWAWSVLGDERS

QFCRVRELRLMDTEKQSRADORITRESQVVELAVKHKAILALQALKEQKLEK

AESLSDKXNDELKXGHLNENRSLQQLKETERELKORLEKQKLEQKQDLQKHLF

RLTQGLQALDRADLLKTERDUEYQLENTQVLYSHEKVKWEGTISQTKLIDFLQAK

MDQPAKKKQVPLQYNELKJALEKAKCAELEALQKTRIELRSAREEAHRAKATDP

HPSTPAQACQIAMSIVRSEPHQPSAMSLAPPSSRRKESSTPEEFSRLKERMHN

IPIHFNVGLNMRATKCAVCLDTVHFQSLGKMCLEQVNMCHPKGKSTCLCPATCGLP

THTEAFRCRDMNSPGLSKSPGSLHLEGMKVPNNRKRGGQGDQSKYIVLEGSKVL

LYDNEAEAGORPVEEFELCLPDGVSIGHAVGASSELANTAKADVPVILKMHSPHT

CWPGRTLYLLAPFPDKQKRWVTALESVAVAGRSREKAEADAKLGNLSLKLEGGDRL

DMNCTLPFSDQVVLVTEBGLYALNVLKNSLTHIPGIGAVFYQIILKQLEKLMIAE

ERALCLVDVKKVKSQAQSLHAPQDVSFNI FEAVKGCFLFAAGKINSJCICAAMS

KVILRYNDLNSKYICKEITSEPCSIHTFYNSILIGIKNFYEDIMQYITLDEFD

KNDHSLAPAVFASNSFPVSIQVANSAGOREBYLLCFHEFVFDVSGYRRSRDLDK

WSRLAPAVREPYLFVTHFNSLEVIQIARSLSGPARVLEI PNRYLGLPAISSGA

LYLASSYQDKLRLVICCKGNLVKESGTEHVRPSTSSPNKRGPTTNEHITREVAS

PAPPEGSHPREPSTPHRYDRRETELRDKSPGRFLEREKSPGRMLSTRERSPOR

LFEDSSRGRUPAGAVRTPLSQVNVKWDQSSV"

ORIGIN

Query Match 54.7%; Score 3593.8; DB 10; Length 4967;

Best Local Similarity 84.1%; Pred. No. 0;

Matches 4249; Conservative 0; Mismatches 527; Indels 279; Gaps 6;

Qy 1136 CCCCTTGGTCCACCTCAAGTCTGACGATGACACCTCAATTTTGTGATGAACAGAGA 1195

Db 1 CCCCTTGGTCCACCTCAAGTCTGACGATGACACCTCAATTTTGTGATGAACAGAGA 60

Qy 1196 AGAATTCGTGGGTTTCATCTCTCCGTGCGAGCTGAGCCCTCAGGCTTCTCGGGTGAAG 1255

Db 61 AGAATTCGTGGGTTTCATCTCTCTGTCGACAGTGGCCCTCGGGCTTCTCAGGCGAG 120

Qy 1256 AATGCGGTTTGTGGGTTTTCGTACAGCAAGGCACTGGGGATTTCTGGTAGATCTGAGT 1315

Db 121 AGTTCGCGTTTGTGGGATTTTCGTACAGCAAGGCACTGGGGATTTCTGGTAGATCTGAGT 180

Qy 1316 CTGTGTGTGGGCTGTGAGCTCCCTCCCAAGACTAGCTCCATGGAAGAAAGAACTCTCA 1375

Db 181 CTGTGTGTGGGCTGTGAGCTCCCTCCCAAGGTTAGTCCATGGAAGAAAGAACTCTCA 240

Qy 1376 TCAAAAGCAAGAGCTCAAGACTCTCAGGACAGTGTCAACAGATGGAGCAGGAAATGA 1435

Db 241 TCAAAAGCAAGAGCTCAAGACTCCAGGACAGTGTCAACAGATGGAGCAGGAAATGA 300

Qy 1436 CCCGTTACATCGAGAGTGTGAGGTTGAGGCTGTGTTAGTTCAGAGAGAGTGGAGC 1495

Db 301 CCCGTTACATCGAGAGTGTGAGGTTGAGGCTGTGTTAGTTCAGAGAGAGTGGAGC 360

Qy 1496 TGAAGGCTCTGAGACTCAGAGATCCCTCTGGAGCAGGACCTTGTCTACCTACATCAG 1555

Db 361 TGAAGGCTCTGAGACTCAGAGATCCCTCTGGAGCAGGACCTTGTCTACCTACATCAG 420

Qy 1556 AATGAGTGTAAAGCGAAGTTTGGAGCAAGCAGGATGGAGGTTGCCAGGAGGATG 1615

Db 421 AATGAGTGTAAAGCGAAGTTTGGAGCAAGCAGGATGGAGGTTGCCAGGAGGATG 480

Qy 1616 ACAAGCACTGACGCTTCTCCATGATATCAGAGCAGGACGCGGAGCTCCAGAGATCA 1675

Db 481 ACAAGCTCTGACGCTTCTCCACGACATCCGAGAGCAGAGCCGGAAGCTCCAGGAGATCA 540

Qy 1676 AAGAGCAGGAGTACAGGCTCAAGTGGAGAAAGAAATGAGTGTGATGAATCAGTTGGAAG 1735

Db 541 AAGAGCAGGAGTACAGGCTCAGGTGGAGGAGATGAGCTGATGAATCAGTTGGAAG 600

Qy 1736 AGGATCTTGTCTAGAGAAAGACGAGTATCTCTACAGATCTGACGTGAGAGAGTCTC 1795

Db 601 AAGACCTGGTGTGACCCCGCAGACGACGAGCTCTCTACGAGTCTGAGCTGAGGAGTCTC 660

Qy 1796 GGCTTGTCTGCTGAAGAAATTCAGCGGAAAGCGACAGAAATGTGACATAAATCTGTGAAG 1855

Db 661 GGCTTGTCTGCTGAGAAATTCAGCGGAAAGCGACAGAAATGTGACATAAATCTGTGAAG 720

QY 1856 CTAAGGATCAAGGAGAGCTGAAGTGGAGAAATATGCGAAACTGGAGAGATCAATGCTG 1915
Db |||||
QY 721 CTAAGGACCAAGGAGAGCTGAAGTGGAGAAATATCCAAACTGGAGAGATCAATGCTG 780
Db |||||
QY 1916 AGCAGCAGCTCAAAATTCAGGAGCTCCAAAGAGAACTGGAGAGAGCTGCAAGAGCGAG 1975
Db |||||
QY 781 AGCAGCAGCTCAAGATCCAGGAGCTCAGAGAGAGCTGGAGAGAGCTGTAAAAGCCAGCA 840
Db |||||
QY 1976 CCGAGAGGAGCTGGAGAGCTCGAGAACCGAGAGAGATTTCTGTAGAGGCATCAGAAAGA 2035
Db |||||
QY 841 CAGAGGCCACCGAGCTCTCGAGAAATATCCGCCAGG-----CAAAGG 882
Db |||||
QY 2036 AGCTGTGGAGCTGAGAGAACCGCGCCATTCCTGGAGAGACAAAGTAAAGAGACTAGAGA 2095
Db |||||
QY 883 AGCAGCAGAGCGGGAG-----CTAGAGAGCTTACAAACCGGGAAGACT 927
Db |||||
QY 2096 CCATGGAGCCTAGAGAAACAGACTGAGAGATGACATCCAGACAAAATCCCAACAGATCC 2155
Db |||||
QY 928 CCTCCGAGGGCATCAAAAAGAGCTGGTGAAGCC----- 962
Db |||||
QY 2156 AGCAGATGGCTGATAAAATTCCTGGAGCTCGAAGAGAACATCGGAGGCCCAAGTCTCAG 2215
Db |||||
QY 963 -----GAGAGCTGGAGAGAAACATCGGAGGCTCAGGTCTCAG 1002
Db |||||
QY 2216 CCCAGACCTAGAGTGCACCTGAAAAGAGAGCAGCACTATGAGGAAAGATTAAG 2275
Db |||||
QY 1003 CTCAACATCTAGAGTACACTTGAACAGAGAGCAAGCACTACGAGGAAAGATCAAGG 1062
Db |||||
QY 2276 TGTGGCAATCAGATTAAGAAAGACTGCTGTACAAAGAGACACTGGAGAACATGATGC 2335
Db |||||
QY 1063 TATTGGACAATCAGATTAAGAAAGACCTGCCACAGAGAGAGCTGGAGAACATGATGC 1122
Db |||||
QY 2336 AGAGACACGAGGAGGAGGCCCATGAGAAAGGCAAAATTTCTCAGCGAACAGAGGCCATGA 2395
Db |||||
QY 1123 AGAGACACGAGGAGGAGGCCCAAGAGAGGCAAGATCTCAGCGAGCAGAGGCCATGA 1182
Db |||||
QY 2396 TCAATGTATGGATTCAGAGATCAGATCCCTGTGAAACAGAGGATTTGGAACTGTCTGAAG 2455
Db |||||
QY 1183 TCAACGGCTGGATTCAGATCCGATCCCTGGAGCAGAGGATCGTGGAGCTGTCGGAAG 1242
Db |||||
QY 2456 CCAATRAACTGGAGCAATAGCAGTCTTTTACCAAGGACATGAGGCCCAAGAG 2515
Db |||||
QY 1243 CCAACAAGCTTGGGCAACAGCAGTCTTTCACCAAGAGAAACATGAGAGGCCCAAGAG 1302
Db |||||
QY 2516 AGATGATTTCTGAACCTCAGCAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGG 2575
Db |||||
QY 1303 AGATGATCTCAGAACTCAGGACGAGAAATTTTACCTGGAGACGAGGCGGGAAGCTGG 1362
Db |||||
QY 2576 AGCCCGACAGCAGAACTGGAGGACGCTGGAGAGATCAGCCACCAAGACCAAGTG 2635
Db |||||
QY 1363 AGGCCAGAACCCAAAGCTGGAAGCAACTGGAGAGATCAGCCACCAAGATCAGATG 1422
Db |||||
QY 2636 ACAAGATCGGCTGCTGGAACCTGGAGACAAAGTTGGCGGAGGTCAGTCTAGAGCAGAGG 2695
Db |||||
QY 1423 ACAAGAGCTCGCTGAGCTGGAACCAAGCTGAGGAGGTCAGCTTGGAGCAGAGG 1482
Db |||||
QY 2696 AGCAGAACTGGAGCTCAAGCGCAGCTGACAGAGCTACAGTCTCCCTGGAGAGCGCG 2755
Db |||||
QY 1483 AGCAGAGCTGGAGCTGGAAGCGCAGCTCAGGAGCTGAGCTGTCCTCGAGGAGCGCG 1542
Db |||||
QY 2756 AGTCACAGTTGACAGCCCTCGAGCTGCAAGCGCGCCCTGGAGAGCAGCTTCGCCAGG 2815
Db |||||
QY 1543 AGTCCAGCTGACGGCCCTGCAAGCGCCGCGGACGCTGGAGAGCAGCTCCGCCAGG 1602
Db |||||
QY 2816 CGAAGACAGAGCTGGAGAGACCAAGAGAGAGCTGAAAGAGAGATCCAGGCACTCAGCG 2875
Db |||||
QY 1603 CGAAGACAGAGCTGGAGAGACCAACCGGAAACCGGAGGAGAGATCCAGGCGCTCAGCG 1662
Db |||||
QY 2876 CACATGAGATGAAATCCAGCGCAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCA 2935
Db |||||
QY 1663 CACATCGGATGAGATCCAGCGCAATTCGATGCCCTTGGCAACAGCTGCACGGTATCA 1722
Db |||||
QY 2936 CAGACCTGGAGGAGCAGCTTAAACCAAGCTACGAGGACAAACGCTGAATCTAAACAAACAAA 2995

Db |||||
QY 1723 CCGACTGGAGAGCAGCTGAACCAAGCTCACCGAGGACAAACCGCAGCTCAACAAACAAA 1782
Db |||||
QY 2996 ACTTCTACTTTGTCAAACAACTCCGATGAGGCTTTCTGGCGCCAAACGACGAGATTTGTACAA 3055
Db |||||
QY 1783 ACTTCTACTTTGTCAAACAACTCCGATGAGGCTTCGGGGCCAAATGACGAGATTTGTCAAG 1842
Db |||||
QY 3056 TCGAAGTGAAGTGAACCATCTCCGCCGGAGATCACGGAACGAGAGATGAGCTTACCA 3115
Db |||||
QY 1843 TCGGAGTGAAGTGAACCATCTCCGCCGTGAGATCACGAGCGGAGATGAGCTCACC 1902
Db |||||
QY 3116 GCCAAGCAAAACGATGAGGCTCTGAAGACCACTGCAACCATGCTGGAGAAACAGGTCA 3175
Db |||||
QY 1903 GCCAGAGCAAAACGATGAGGCTCTGAAGACGACATGACCATGTTGGAGAGCAGGTCC 1962
Db |||||
QY 3176 TGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAAAAGAGCGGAGTGGAGGCTTGA 3235
Db |||||
QY 1963 TGGACTGGAGGCCCTGAAACGACGAGCTGTGGAGAGAGCGGAGCTGGGAGGCTTGGC 2022
Db |||||
QY 3236 GGAGCGTCTCGGTGATGAGAAATCCCAAGTTTGTAGTGTGGGTTCGAGAGCTGCAGAGAA 3295
Db |||||
QY 2023 GGAGCGTCTTGGCGACGAGAACTCCAGTTTGGAGTTCGAGTTCGAGAGTACAGAGGA 2082
Db |||||
QY 3296 TGTGGACACCGAGAAACAGAGCAGGCGAGGCGGATCAGCGGATCACCGAGTCTGCC 3355
Db |||||
QY 2083 TGTGGACACCGAGAAACAGAGCAGGCGGATGAGGCGGATCAGCGGATCACCGAGTCTGCC 2142
Db |||||
QY 3356 AGTGTGCTGCTGGAGTGAAGGAGCAAGGCTGAGATTCGCTCTGCTGAGAGGCTC 3415
Db |||||
QY 2143 AGTGTGCTGAGTGGCGTGAAGGAAACAAAGGCGGAGATTCCTGCTCTGAGAGGCTC 2202
Db |||||
QY 3416 TCAAGACAGCAAGCTGAAGGCCGAGAGCCCTCTCTGCAAGCTCAATGACCTGGAGAGA 3475
Db |||||
QY 2203 TCAAGACAGCAAGCTCAAGCCGAGAGCCCTGTCGACNAAGCTCAACGCTGAGAGGA 2262
Db |||||
QY 3476 AGCATGCTATCTTGAATGAATGCCGAGGCTTACAGCAGAGCTGGAGAGCTGGAAG 3535
Db |||||
QY 2263 AGCAGCGCATCTGGAGATGAACCGCGGAGCTTACAGCAGAAACTAGAGACAGAGCGG 2322
Db |||||
QY 3536 AGCTCAACACAGAGCTTCTGGAAGCAAGCAAAATTTACAGCAGCAGATGAGCTGAGAG 3595
Db |||||
QY 2323 AGCTCAACACAGAGCTTCTGGAGAGCAAGCAAAATTTGAGCAGCAGATGAGCTGAGAG 2382
Db |||||
QY 3596 AAAATCAATTTCCGCTGTGACTCAAGGATGCAAGAGCTCTAGATCGGCTGATCTAC 3655
Db |||||
QY 2383 AGAACCAATCTTCAGACTGACGCAAGGCTGAGAGGCGCTGGACCGGCGGATCTGC 2442
Db |||||
QY 3656 TGAAGACAGAAAGAGTGAATGGAGTATCAGCTGGAAACATTCAGGTTCTCTATTCTC 3715
Db |||||
QY 2443 TGAAGACAGAAAGAGAGGACCTGGAGTACGAGTGGAAACATTCAGGTTCTCTACTCTC 2502
Db |||||
QY 3716 ATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGATTTCTGC 3775
Db |||||
QY 2503 ACAGAGAAAGTGAATAATGGAAGGCACTCTCTCAGCAAAACCAAACTCATTTGATTTCTGC 2562
Db |||||
QY 3776 AAGCCAAATGGAACCACTCTTAAGAGAAAGAGGTTTATTTAGTCGAGGAAAGAGG 3835
Db |||||
QY 2563 AAGCCAAATGGAACGAGCTCTTAAAGAAAA----- 2595
Db |||||
QY 3836 ACCCTGCTTTACCCACACAGGTTCTCTGAGTACAATGAGCTGGAAGCTGGCCCTGGAGA 3895
Db |||||
QY 2596 -----AGGTTCTCTGAGTACAATGAGTGAAGCTAGCCCTGGA 2637
Db |||||
QY 3896 AGGAGAAAGCTCGTGTGAGAGCTAGAGAGAGCCCTTCAGAGAGCCGAGTGGAGTCC 3955
Db |||||
QY 2638 AGGAGAAAGCCGATGCGGAGCTGGAGAGGCGCTTCAGAGAGCCCGCATGGAATCC 2697
Db |||||
QY 3956 GGTCCGCGCGGAGAGAGCTGCCCAACCGCAAGCAACGAGCAGCCCAACCCATCCACGC 4015
Db |||||
QY 2698 GCTCTGCGCGGAGAGAGTGGCCCAACCGCAAGCAACGAGCAGCCCGCATCCGCTACGC 2757
Db |||||
QY 4016 CAGCCACGCGGAGGAGAGATGCGCATGTCCGCTGCGGTGCGCAGAGCAGCAGC 4075
Db |||||

2758 CAGCCACTGCGAGCAGACAGATCGCCATGCTGCCATTGTGGCTGCGCGGACGACCAGC 2817
4076 CCAAGTCCCATGAGCTGCTGGCCCGCCATCCAGCCGCGAAGAGGAGTCTTCAACTCCAG 4135
2818 CAGTGCCTATGAGCTGCTGGCCCGCCATCCAGCCGCGAAGAGGAGTCTTCAACTCCAG 2877
4136 AGGAATTTAGTGGCGCTCTTAAGGAACGAGTGCACCAACATATTCCTCACCGGATTCAGG 4195
2878 AGGAATTCAGCGCGCTCTGAAGAGCGCATGCACCAACATCCCTCACCGGTTAATG 2937
4196 TAGGACTGAACATCGAGCCACAAAGTGTCTGTGTCTGTGATACCGTGCACCTTGGAC 4255
2938 TGGGCTGAACATGAGAGCCACAAAGTGGCGGTGTGTGTGATCTGTGACCTTGGAC 2997
4256 GCCAGCATCCAAATGCTCGAATGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 4315
2998 GCCAGCATCCAAATGCTCGAATGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 3057
4316 TGGCAGCATCTGCGGCTTGCCTGTCTGATATGCGACACACTTCCACCGAGGCTTCTGCC 4375
3058 TGGCGCTACTGTGGCTGCCAGCTGAATATGCCACACACTTCACTGAGGCTTCTGCC 3117
4376 GTGACAAAATGAATCCCAAGTCTCCAGCAAGAGGCCAGCAGCTTGCACCTGG 4435
3118 GGGACAAAATGAATCTCCCGGGCTCCAGCAAGAGGCTGCGAGCAGCTTGCACCTGG 3177
4436 AAGGTGGATGAAGTGGCCAGGAATACAAACGAGCAGCAAGGCTGGGACAGGAAGT 4495
3178 AAGGTGGATGAAGTGGCCAGGAATACAAACGAGCAGCAAGGCTGGGACAGGAAGT 3237
4496 ACATTTCTGGAGGATCAAAAGTCTCATTTATGACAAATGAAGCAGAGGAGCTGGAC 4555
3238 ACATTTCTGGAGGATCAAAAGTCTCATTTATGACAAATGAAGCAGAGGAGCTGGAC 3297
4556 AAGGCGCGTGGAAAGTGTGAGTGTGCTTCCGAGCGGATGTATCTATTCTATGGTG 4615
3298 AAGGCGCGTGGAAAGTGTGAGTGTGCTTCCGAGCGGATGTATCTATTCTATGGTG 3357
4616 CGTTGGTGGTTCGAACTCGCAATACAGCCAAAGCA----- 4653
3358 CGTTGGTGGTTCGAACTTCGAAATACCGCCAAAGCAGATGTCCCATATCCTCGAAG 3417
4654 ----- 4653
3418 TGGAGTCTCACCCACACACACCTGCTGGCTGGGAGAACCTCTACTTGTAGCACCCA 3477
4654 ----- 4653
3478 GCTTCCCGCAGCAGCGCTGGGTCAACCGCTTAGAATCTGTGTGGCGAGGAGAG 3537
4654 -----GAAAGCAGAGCTGATCTAAACTGCTTGGAACTCCCTGCTGAACTGG 4705
3538 TTTCTAGGGAAGGCGGAGCGAGTGTCTAAATTAATTTAGTGAATCTCTGTGAAACTGG 3597
4706 AAGTGATGACCGCTTAGACATGAATGACGCTGCTTCAAGTCAAGTGGTGGTGGTGG 4765
3598 AGGCGATGACCGGCTTGACATGAATGACCTGCTGCTTCAAGTCAAGTGGTGGTGGTGG 3657
4766 TGGGCAACGAGGAGGCTCTAGCGCTGAAATGCTTTGAAAGTCTTCAAGTGGTGGTGGTGG 4825
3658 TGGGCAACGAGGAGGCTCTAGCGCTGAAATGCTTTGAAAGTCTTCAAGTGGTGGTGGTGG 3717
4826 CAGGAATTTGAGCAGTCTTCAAAATTAATTAATCAAGGAGCTGGGAGAGTCTCATGA 4885
3718 CAGGATTTGCGCAGTCTTCAAAATTAATTAATCAAGGAGCTGGGAGAGTCTCATGA 3777
4886 TAGCAGGAGAGGCGGCTCTGTGTGTGGAGCTGAAGAGTGAAGTGAAGTGAAGTGAAGTGA 4945
3778 TAGCAGGAGAGGCGGCTCTGTGTGTGGAGCTGAAGAGTGAAGTGAAGTGAAGTGAAGTGA 3837
4946 CCAAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5005
3838 CGCAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3897

5006 GCTGCCACTTTTGGGCGCAGGCAAGATTGAGAACGGGCTCTGCATCTGTGACGCCATGC 5065
3898 GCTGCCACTTTTGGGCGCAGGCAAGATTGAGAACGGGCTCTGCATCTGTGACGCCATGC 3957
5066 CCAGCAAGTGTCTCATTTCTCGCTTACAAAGAAACCTCAGCAAAATCTGATCCGGAAG 5125
3958 CAAGCAAGTGTCTCATTTCTCGCTTACAAAGAAACCTCAGCAAAATCTGATCCGGAAG 4017
5126 AGATAGAGCCTCAGAGCCTGAGCTGTATCCACTTCAAAATCAGATCTCTCATTTG 5185
4018 AGATAGAGCCTCAGAGCCTGAGCTGTATCCACTTCAAAATCAGATCTCTCATTTG 4077
5186 GAACCAATAAATTTACGAAATGACATGAAGCAGTACAGCTCAGGAAATTCCTGGATA 5245
4078 GGACCAATAAATTTATGAGATGACATGAAGCAGTACAGCTTGTATGAGTTCCTGGACA 4137
5246 AGAATGACCAATTCCTTGGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5305
4138 AGAAGCAGCAATTCCTTGGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4197
5306 CAATCGTGCAGGTGAACAGCGCAGGCGAGAGAGTACTTGTGTGTGTGTGTGTGTGTGTGT 5365
4198 CCATTTGTGAGCGCAACAGCGCGGCGAGAGAGTACTTGTGTGTGTGTGTGTGTGTGTGT 4257
5366 TTGAGT 5425
4258 TTGAGT 4317
5426 GCTTACCTTTGGCTTGT 5485
4318 GCTTACCTTTGGCTTGT 4377
5486 TCGAAGTAATGAGATCCAGCAGCTTCTCAGCAGGAGCCCTGCGCGAGCGTACCTGG 5545
4378 TGAAGTCAATGAGATCCAGCAGATCTTCTCAGGAGAGCTTCTCAGGAGAGTACCTGG 4437
5546 ACATCCCGAAGCCCGCTTCTCAGGAGAGCTTCTCAGGAGAGTACCTTCTCAGGAGAGT 5605
4438 AAATTCGAAGCCCTGCTTCTCAGGAGAGCTTCTCAGGAGAGTACCTTCTCAGGAGAGT 4497
5606 CCTCATACCGAGTAAATTAAGGCTCATTTGCTCAAGGAGAACTCTGCTGAAGAGTCCG 5665
4498 CCTCATACCGAGTAAATTAAGGCTCATTTGCTCAAGGAGAACTCTGCTGAAGAGTCCG 4557
5666 GCACTGAACACCGCGGCGCTTCTCAGGAGAGCTTCTCAGGAGAGCTTCTCAGGAGAGCT 5725
4558 GCACTGAACACCGCGGCGCTTCTCAGGAGAGCTTCTCAGGAGAGCTTCTCAGGAGAGCT 4617
5726 CCAAGTCAACGAGCAGATCAACGAGCGTGTGCTTCAAGGAGAACTCTGCTGAAGAGTCCG 5785
4618 CAAATACAGCAGCAGATCAACGAGCGTGTGCTTCAAGGAGAACTCTGCTGAAGAGTCCG 4677
5786 CCAAGTCAACGAGCAGATCAACGAGCGTGTGCTTCAAGGAGAACTCTGCTGAAGAGTCCG 5839
4678 CCAAGTCAACGAGCAGATCAACGAGCGTGTGCTTCAAGGAGAACTCTGCTGAAGAGTCCG 4737
5840 AGCTGCAAGGAGCAGTCTTCTGCGCGCTTCTGAGCGAGAGTCTTCTGAGCGAGAGTCTTCT 5899
4738 AGCTGCAAGGAGCAGTCTTCTGCGCGCTTCTGAGCGAGAGTCTTCTGAGCGAGAGTCTTCT 4797
5900 TGCTCAGCAGCGCGAGAGCGGTCTTCTGAGCGAGAGTCTTCTGAGCGAGAGTCTTCTGAGCG 5959
4798 TGCTCAGCAGCGCGAGAGCGGTCTTCTGAGCGAGAGTCTTCTGAGCGAGAGTCTTCTGAGCG 4857
5960 GGCTGCTGCGGAGCGGTGAGGAGCTTCTGAGCGAGAGTCTTCTGAGCGAGAGTCTTCTGAGCG 6019
4858 GGCTGCTGCGGAGCGGTGAGGAGCTTCTGAGCGAGAGTCTTCTGAGCGAGAGTCTTCTGAGCG 4917
6020 GTGCTCTCAAGTTT 6034
4918 CTTTCAATTAAGTCT 4932

Mon Jul 26 15:37:17 2004

us-10-017-216-1.rge

Page 48

Search completed: July 4, 2004, 03:42:04
Job time : 24322 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2004, 09:03:14 ; Search time 396 Seconds
(without alignments)
9212.745 Million cell updates/sec

Title: US-10-017-216-1
Perfect score: 6574
Sequence: 1 agagcgccagtgaggagat.....atcgagaatgagtttaga 6574

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PTCUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1397.8	21.3	1515	4	US-09-804-471A-1
2	1397.8	21.3	1515	4	US-10-238-709-1
3	768	11.7	1133	4	US-09-916-204-1
4	258	3.9	258	4	US-09-016-434-513
5	220.8	3.4	2726	2	US-08-422-699A-12
6	220.8	3.4	2726	2	US-08-422-705B-12
7	214	3.3	2511	2	US-08-422-699A-8
8	214	3.3	2511	2	US-08-422-705B-8
9	213.6	3.2	3182	1	US-08-484-044-11
10	205	3.1	174493	4	US-09-804-471A-3
11	205	3.1	174493	4	US-10-238-709-3
12	189.2	2.9	2706	2	US-08-630-822A-61
13	189.2	2.9	2706	2	US-09-005-069-61
14	189.2	2.9	2706	4	US-09-171-158A-20
15	189.2	2.9	2706	4	US-09-004-730A-20
16	189.2	2.9	2706	4	US-08-981-799A-20
17	181.2	2.8	4363	4	US-08-685-576-5
18	179.6	2.7	4848	4	US-09-976-594-295
19	178.8	2.7	48763	4	US-09-916-204-3
20	172.2	2.6	4065	4	US-09-016-434-1105
21	172.2	2.6	4739	3	US-08-685-871-1
22	168.4	2.6	5053	2	US-08-685-576-2
23	150.2	2.3	3323	2	US-08-422-699A-10
24	150.2	2.3	3323	2	US-08-422-706B-10
25	122.6	1.9	3155	4	US-09-442-100-7
26	122.6	1.9	3155	4	US-08-939-106-7
27	122.6	1.9	3155	4	US-09-442-102-7

28	122.2	1.9	1498	4	US-09-509-902A-6	Sequence 6, Appli
29	122.2	1.9	1961	4	US-09-509-902A-15	Sequence 15, Appli
30	122.2	1.9	5276	4	US-09-233-857-2	Sequence 2, Appli
31	117.8	1.8	3018	2	US-08-850-150-6	Sequence 6, Appli
32	117.8	1.8	3018	3	US-09-338-132-6	Sequence 6, Appli
33	111.4	1.7	1935	2	US-08-878-989-11	Sequence 11, Appli
34	111.4	1.7	1935	3	US-09-272-796-11	Sequence 11, Appli
35	109.8	1.7	3213	4	US-09-442-100-5	Sequence 5, Appli
36	109.8	1.7	3213	4	US-08-939-106-5	Sequence 5, Appli
37	109.8	1.7	3213	4	US-09-442-102-5	Sequence 5, Appli
38	105	1.6	3489	2	US-08-728-223A-1	Sequence 1, Appli
39	105	1.6	3489	4	US-09-298-568-1	Sequence 1, Appli
40	105	1.6	3489	4	US-09-410-399-1	Sequence 1, Appli
C 41	105	1.6	32207	2	US-08-770-379-20	Sequence 20, Appli
C 42	105	1.6	32207	3	US-08-757-669A-20	Sequence 20, Appli
C 43	105	1.6	32207	4	US-09-230-371A-20	Sequence 20, Appli
44	104.8	1.6	2160	3	US-09-588-256-1	Sequence 1, Appli
45	101.4	1.5	5720	4	US-09-442-100-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-804-471A-1
; Sequence 1, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Human
US-09-804-471A-1

Query Match		21.3%	Score 1397.8;	DB 4;	Length 1515;
Best Local Similarity		99.5%	Pred. No. 0;		
Matches 1402;		Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	13	GGGAGATGTTGAAGTTCAAATATGGAGCGCGGAATCTTTTGGATGCTGCTGCTGAA	72		
Db	1	GGGAGATGTTGAAGTTCAAATATGGAGCGCGGAATCTTTTGGATGCTGCTGCTGAA	60		
Qy	73	CCCATTCGATGAGTGGGGCTCCAGCTGAATCTCTTCCAGGGGAAACACCTTTATG	132		
Db	61	CCCATTCGATGAGTGGGGCTCCAGCTGAATCTCTTCCAGGGGAAACACCTTTATG	120		
Qy	133	ACTCAACAGAGATGCTCTCTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTC	192		
Db	121	ACTCAACAGAGATGCTCTCTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTC	180		
Qy	193	TTTGAAGATGCGATCAGCTGCTCTGATGAAGATTAAGCACTGAGCACTTTCGG	252		
Db	181	TTTGAAGATGCGATCAGCTGCTCTGATGAAGATTAAGCACTGAGCACTTTCGG	240		
Qy	253	AAGTATTCGCACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGAAAGGACTTCGAA	312		
Db	241	AAGTATTCGCACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGAAAGGACTTCGAA	300		
Qy	313	GTGAGAAGTCTTGTAGTTGTGCTACTTTGCTGAAGTGCAGTGCTGTAAGAGAAGCA	372		
Db	301	GTGAGAAGTCTTGTAGTTGTGCTACTTTGCTGAAGTGCAGTGCTGTAAGAGAAGCA	360		
Qy	373	ACCGGGACATCTCTGATGAAGTGAAGCAAGCAAGCTTTATGCGCCAGGAGCAG	432		

```
Db 361 ACCGGGACATCTATGCTATGAAAGTGATGAAGAAAGGCTTTATTGGCCCGCAGGAGCAG 420
Qy 433 GTTTCATTTTTCAGGAAGAGCGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCG 492
Db 421 GTTTCATTTTTCAGGAAGAGCGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCG 480
Qy 493 CAATTACAGTATGCCCTTCAGGACAAATACCTTTATCTGATGAGGAATATCAGCCT 552
Db 481 CAATTACAGTATGCCCTTCAGGACAAATACCTTTATCTGATGAGGAATATCAGCCT 540
Qy 553 GGAGGGACCTGCTGTCACCTTTTGAATAGATAGAGGACAGTATGATAAAACCTGATA 612
Db 541 GGAGGGACCTGCTGTCACCTTTTGAATAGATAGAGGACAGTATGATAAAACCTGATA 600
Qy 613 CAGTTTACCTAGCTAGCTGATTTGGCTGTTTCAGAGGTTTCATCTGATGGATACGTG 672
Db 601 CAGTTTACCTAGCTAGCTGATTTGGCTGTTTCAGAGGTTTCATCTGATGGATACGTG 660
Qy 673 CATCGAGACATCAAGCCTGAGAACTTCTCGTTGACCGCACAGGACACATCAAGCTGGTG 732
Db 661 CATCGAGACATCAAGCCTGAGAACTTCTCGTTGACCGCACAGGACACATCAAGCTGGTG 720
Qy 733 GATTTTGGATCTCCCGGAAATGAAATTCAAACAGATGTTGAATGCCAAATCCCGATT 792
Db 721 GATTTTGGATCTCCCGGAAATGAAATTCAAACAGATGTTGAATGCCAAATCCCGATT 780
Qy 793 GGGACCCAGATACATGCTGCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGC 852
Db 781 GGGACCCAGATACATGCTGCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGC 840
Qy 853 ACCTACGGCTGACTGTGACTGCTGCTGAGTGGGCTGATGCTGATGAGATGATTTAT 912
Db 841 ACCTACGGCTGACTGTGACTGCTGCTGAGTGGGCTGATGCTGATGAGATGATTTAT 900
Qy 913 GGGAGATCCCTTTCGAGAGGAACTCTGCGCAGAACTTCAATAACATTTATGATTTTC 972
Db 901 GGGAGATCCCTTTCGAGAGGAACTCTGCGCAGAACTTCAATAACATTTATGATTTTC 960
Qy 973 CAGCGTTTTCGAAATTCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1032
Db 961 CAGCGTTTTCGAAATTCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
Qy 1033 CAAAGCTTGTGTCGCGCAGAAAGAGAGACTGAAGTTCGAGGCTTTGCTGCTGCTGCT 1092
Db 1021 CAAAGCTTGTGTCGCGCAGAAAGAGAGACTGAAGTTCGAGGCTTTGCTGCTGCTGCT 1080
Qy 1093 TTCTTCTCTAAATTCGATGGAACAACTTCGTAACCTCTCTCTCTCTCTCTCTCTCTCT 1152
Db 1081 TTCTTCTCTAAATTCGATGGAACAACTTCGTAACCTCTCTCTCTCTCTCTCTCTCTCT 1140
Qy 1153 CTCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1212
Db 1141 CTCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1200
Qy 1213 TCCTCTCCGTCGAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTGTGGGG 1272
Db 1201 TCCTCTCCGTCGAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTGTGGGG 1260
Qy 1273 TTTTCTGACAGAGGACCTGCGGATTTCTTGTGTGATCTGAGTCTGTTGTGCGGTCTG 1332
Db 1261 TTTTCTGACAGAGGACCTGCGGATTTCTTGTGTGATCTGAGTCTGTTGTGCGGTCTG 1320
Qy 1333 GACTTCCCTCCCAAGACTAGCTCCATGGAAGAACTCTCTCATCAAAAGCAAGAGCTTA 1392
Db 1321 GACTTCCCTCCCAAGACTAGCTCCATGGAAGAACTCTCTCATCAAAAGCAAGAGCTTA 1380
Qy 1393 CAAAGCTCTCAGGACAGTGTCAAGAT 1421
Db 1381 CAAAGCTCTCAGGACAGTGTCAAGAT 1409
```

RESULT 2

US-10-238-709-1

```
; Sequence 1, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Human
; US-10-238-709-1

Query Match 21.3%; Score 1397.8; DB 4; Length 1515;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1402; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 13 GGGGAGATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAA 72
Db 1 GGGGAGATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAA 60
Qy 73 CCCATTCCAGCCGGGCTCCAGCTGAATCTGTTCTTCAGGGGAAACACACCTTTATG 132
Db 61 CCCATTCCAGCCGGGCTCCAGCTGAATCTGTTCTTCAGGGGAAACACACCTTTATG 120
Qy 133 ACTCAACAGCAGATGCTCTCTCTTTCCGAGAGAGGATATTAGATGCTCTTTGTTCTC 192
Db 121 ACTCAACAGCAGATGCTCTCTCTTTCCGAGAGAGGATATTAGATGCTCTTTGTTCTC 180
Qy 193 TTTCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
Db 181 TTTCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 253 AAGTATTCGACACCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
Db 241 AAGTATTCGACACCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 313 GTCAGAACTCTTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
Db 301 GTCAGAACTCTTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 373 ACCGGGACATCTATGCTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 432
Db 361 ACCGGGACATCTATGCTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
Qy 433 GTTTCATTTTTCAGGAAAGCGGAACTATTTCTCGAAGCAACAGCCCTGCTGCTGCTGCTG 492
Db 421 GTTTCATTTTTCAGGAAAGCGGAACTATTTCTCGAAGCAACAGCCCTGCTGCTGCTGCTG 480
Qy 493 CAATTACAGTATGCTTTCAGGACAAATATCACTTTATCTGATGAGGATATCAGCCT 552
Db 481 CAATTACAGTATGCTTTCAGGACAAATATCACTTTATCTGATGAGGATATCAGCCT 540
Qy 553 GGAGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612
Db 541 GGAGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 613 CAGTTTACCTAGCTAGCTGATTTGGCTGTTTCAGAGGTTTCATCTGATGGATACGTG 672
Db 601 CAGTTTACCTAGCTAGCTGATTTGGCTGTTTCAGAGGTTTCATCTGATGGATACGTG 660
Qy 673 CATCGAGACATCAAGCCTGAGAACTTCTGTTGACCGCACAGGACACATCAAGCTGGTG 732
Db 661 CATCGAGACATCAAGCCTGAGAACTTCTGTTGACCGCACAGGACACATCAAGCTGGTG 720
Qy 733 GATTTTGGATCTGCGCGGAAATGAAATTCAAACAGATGTTGAATGCCAAATCCCGATT 792
Db 721 GATTTTGGATCTGCGCGGAAATGAAATTCAAACAGATGTTGAATGCCAAATCCCGATT 780
```

```

QY 793 GGGACCCAGATTACATGGCTCCTCAAGTGTGACTGTGATGAACGGGGATGGAAAAGGC 852
DB 781 GGGACCCAGATTACATGGCTCCTCAAGTGTGACTGTGATGAACGGGGATGGAAAAGGC 840
QY 853 ACTACGGCTGGAGTGTGACTGTGCTGAGTGGCGGTGATTGCTATGAGATGATTTAT 912
DB 841 ACTACGGCTGGAGTGTGACTGTGCTGAGTGGCGGTGATTGCTATGAGATGATTTAT 900
QY 913 GGGAGATCCCCCTTCGACAGAGGAACCTCTGCGCAAGACCTTCAATAACATTATGAATTC 972
DB 901 GGGAGATCCCCCTTCGACAGAGGAACCTCTGCGCAAGACCTTCAATAACATTATGAATTC 960
QY 973 CAGCGGTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1032
DB 961 CAGCGGTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
QY 1033 CAAAGCTTGTGCGGCGCAGAAAAGAGAGACTGAAGTGTGAAGTGTGAAGTGTGAAGT 1092
DB 1021 CAAAGCTTGTGCGGCGCAGAAAAGAGAGACTGAAGTGTGAAGTGTGAAGTGTGAAGT 1080
QY 1093 TTCTTCTCTAAATTTGACTGGAAACAACTTCCTGCTCCCTCCCTTCGTTCCCAACC 1152
DB 1081 TTCTTCTCTAAATTTGACTGGAAACAACTTCCTGCTCCCTCCCTTCGTTCCCAACC 1140
QY 1153 CTCAGTCTGACGATGACACCTCCCAATTTGATGAACAGAGAGAAATTCGTGGGTTCA 1212
DB 1141 CTCAGTCTGACGATGACACCTCCCAATTTGATGAACAGAGAGAAATTCGTGGGTTCA 1200
QY 1213 TCCTCTCCGTGCCAGCTGAGCCCTCAGCCCTTCTCGGGTGAAGAACTCCCGTTGTGGGG 1272
DB 1201 TCCTCTCCGTGCCAGCTGAGCCCTCAGCCCTTCTCGGGTGAAGAACTCCCGTTGTGGGG 1260
QY 1273 TTTTGTACAGAAAGCACTGGGATTTCTGTAGATCTGAGTCTGTTGTGCGGCTG 1332
DB 1261 TTTTGTACAGAAAGCACTGGGATTTCTGTAGATCTGAGTCTGTTGTGCGGCTG 1320
QY 1333 GACTCCCTGCCAGACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTA 1392
DB 1321 GACTCCCTGCCAGACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTA 1380
QY 1393 CAAGACTCTCAGGACAAAGTGTCAAGAT 1421
DB 1381 CAAGACTCTCAGGACAAAGTGTCAAGAT 1409

RESULT 3
US-09-916-204-1
; Sequence 1, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THERGOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1133
; TYPE: DNA
; ORGANISM: Human
US-09-916-204-1

Query Match 11.7%; Score 768; DB 4; Length 1133;
Best Local Similarity 99.4%; Pred. No. 7.3e-197;
Matches 771; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGCGCGCAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCCTTTGATGCT 60
DB 36 AGAGCGCGCAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCCTTTGATGCT 95

```

```

QY 61 GGTGCTCTGAACCCCAATTGCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAA 120
DB 96 GGTGCTCTGAACCCCAATTGCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAA 155
QY 121 CCACCCCTTTATGACTCAACAGCAGATGTCCTCTTCCCGAGAGGGAATATTAGATGCC 180
DB 156 CCACCCCTTTATGACTCAACAGCAGATGTCCTCTTCCCGAGAGGGAATATTAGATGCC 215
QY 181 CTCTTTGTTCTCTTTGAAGAAATGACGTCAGCTGCTCTGATGAAGATTTAAGCAGTGAGC 240
DB 216 CTCTTTGTTCTCTTTGAAGAAATGACGTCAGCTGCTCTGATGAAGATTTAAGCAGTGAGC 275
QY 241 AACTTTGTCGGGAAGTATTCGGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCA 300
DB 276 AACTTTGTCGGGAAGTATTCGGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCA 335
QY 301 AAGGACTTTCGAAGTCAGAAGTCTTGTAGGTTGTGGTCACTTTGCTGAAGTGCAGGTGGTA 360
DB 336 AAGGACTTTCGAAGTCAGAAGTCTTGTAGGTTGTGGTCACTTTGCTGAAGTGCAGGTGGTA 395
QY 361 AGAGAGAAAGCAACCGGGGACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTG 420
DB 396 AGAGAGAAAGCAACCGGGGACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTG 455
QY 421 GCCCAGGACAGGTTTCATTTTGGAGGAGAGCGGAACATATTATCTCGAAGCACAAAGC 480
DB 456 GCCCAGGACAGGTTTCATTTTGGAGGAGAGCGGAACATATTATCTCGAAGCACAAAGC 515
QY 481 CCGTGGATCCCCCAATTACAGTAGTGCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG 540
DB 516 CCGTGGATCCCCCAATTACAGTAGTGCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG 575
QY 541 GAATATCAGCTCGAGGGGACCTTCTGCTCACTTTTGAATAGTATGAGGACCAAGTTAGAT 600
DB 576 GAATATCAGCTCGAGGGGACCTTCTGCTCACTTTTGAATAGTATGAGGACCAAGTTAGAT 635
QY 601 GAAACCTCTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTTTTCACAGCGTTTCATCTG 660
DB 636 GAAACCTCTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTTTTCACAGCGTTTCATCTG 695
QY 661 ATGGGATACGTGATCGAGACATCAAGCTCGAGAAACATTTCTGTTGACCCGACAGGACAC 720
DB 696 ATGGGATACGTGATCGAGACATCAAGCTCGAGAAACATTTCTGTTGACCCGACAGGACAC 755
QY 721 ATCAAGCTGTGGATTTTGGATCTGCCGCAAAATGAATTTCAACCAAGATGGTGA 776
DB 756 ATCAAGCTGTGGATTTTGGATCTGCCGCAAAATGAATTTCAACCAAGATGGTGA 811

RESULT 4
US-09-016-434-513
; Sequence 513, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:

```


Db 663 TTGTCATGCCCATAGACTCGGTGGTCCACCGGCTTGGCTACGTGCACAGGGACATCAAAACCG 722
Qy 692 AGAACAATCTGTTGACCCACAGACACATCAAGCTGTGTGATTTGGATCTGCCGGA 751
Db 723 ACACATCTGTTGACCGCTGTGGCCACATCGCTGGCCGACTTCGGCTTTCCTCA 782
Qy 752 AAATGAATCAAAAGATGTTGAATGCAAACTCCGATTCGGACCCGACATACATGG 811
Db 783 AGCTGGGGCAGATGGAACGGTGGCTGCTGGTGGCTGGGACCCGACACTACCTGT 842
Qy 812 CTCCTGAAGTGTGA---CTGTGATGAACGGGGATGGAAAGGCACCTACGGCTGGACT 868
Db 843 CCCCCGAGATCTGACAGCTGTGGGGGTGGGCTGGGACAGGAGCTACGGGCCCGAGT 902
Qy 869 GTGACTGTGGTCAAGTGGGCGTGAATTCCTATCAGATGATTTATGGGAGATCCCCCTTCG 928
Db 903 GTGACTGTGGGCGTGGTGTATTCGCTATGAAATGTTCTATGGCAGACGCCCTTCT 962
Qy 929 CAGAGGAACCTTCTGCCAGAACTTCAATAAATATGAATTTCCAGCGGTTTTGAAT 988
Db 963 ACGGGATTCACGGGGGAGACTATGGCAAGATCGTCCACTACAGGAGACCTCTCTC 1022
Qy 989 TTCCAGATCACCCAAAGTGAGCAGT---GACTTCTTGATCTGATTCAAAGCTTGTGT 1045
Db 1023 TGGCGTGTGACGAGGAGGCTCCCTGAGGAGGCTCGAGACTTCAATCAGCGGTTGCTGT 1082
Qy 1046 GGGCCAGAAAAGAGACTGAAGTTTGAAGTCTTTGTGTC-----CATCCTTCTT 1096
Db 1083 GTCCCCGGAGACACGCTGGGCGGGTGGAGCAGGAGCTTCGGGACACATCCCTTCT 1142
Qy 1097 TCTTAATATTAAGTGGAAACATTCGTAATCTCTCTCCCTTCTGTTCCACCCCTCA 1156
Db 1143 TCTTGGCTCTGACTGGGATGTTCTCCGGGACAGGCTGCCCCCTTTACACCGGATTTGG 1202
Qy 1157 AGTCTGACGATGACACTCCAAATTTGA 1184
Db 1203 AAGTCCACCGACACATCAACTTGA 1230

RESULT 6

US-08-422-706B-12
; Sequence 12, Application US/08422706B
; Patent No. 5977333
; GENERAL INFORMATION:
; APPLICANT: Brook, J. David
; APPLICANT: Housman, David E.
; APPLICANT: Shaw, Duncan J.
; APPLICANT: Harley, Helen G.
; APPLICANT: Johnson, Keith J.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
; TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02713
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,706B
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,543
; FILING DATE: 08-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/023,612

; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/839,255
; FILING DATE: 20-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01545
; FILING DATE: 19-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00253
; FILING DATE: 05-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB9202485.0
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-5830A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-422-706B-12
Query Match 3.4%; Score 220.8; DB 2; Length 2726;
Best Local Similarity 53.7%; Pred. No. 6e-49;
Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;
Qy 212 CTGCTCTGATGAAGATTAAAGCAGTGGAGCACTTTGTCGGAAGTATTCGACACCATAG 271
Db 243 CCGAACTGGCCAGGACAAAGTACGTGGCCGACITCTTGCACTGGCGGAGCCCATCTGTG 302
Qy 272 CTGAGTTACAGGAGCTCCAGCCTTCGGCAAAGAGCTTCGAAGTCAGAACTCTTAGTGT 331
Db 303 TGAGGCTTAGGAGGTCGACTGCGAGAGGAGCACTTCGAGATTCTGAAGTGATCGAC 362
Qy 332 GTGTCTACTTTGTGAAGTGCAGGTGTAAGAGAAAGCAACCGGGGACATCTATGCTA 391
Db 363 CGCGGCGGTTTCAAGGAGTACGCGTGTGTAAGATGAAGCAGACGCGGCGGAGTATGCCA 422
Qy 392 TGAAGTGTATGAAGAAGAGCTTTATTGGCCAGGAGCAGGTTTCAATTTTTCAGGAAG 451
Db 423 TGAAGATCATGAACAGTGGGACATGCTGAAGAGGGCGAGGTGCTGCTTCGCTGAGG 482
Qy 452 AGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCGCAATTAAGTATGCTTTTC 511
Db 483 AGAGGAGCTGTTGTTGTAATGGGACCGGCGGTGGATCAGCAGCTGCACCTTCGCTTCC 542
Qy 512 AGGACAAAATCACCTTTATCTGATGAGGAATATCAGCCTGGAGGGGACTTGTGTGAC 571
Db 543 AGGATGAGAACTACTCTGTACCTGTGATGAGTATTACGTGGGCGGAGACTGCTGACAC 602
Qy 572 TTTTGAATAGATATCAGGACCCAGTTAGATGAAAACCTGATACAGTTTTCCTAGCTGAGC 631
Db 603 TGCTGAGCAAGTTTGGGAGCGGATTCGGCCGAGATGGCGCTTCTACCTGGCGGAGA 662
Qy 632 TGAATTTGGCTGTTCACAGCGTTTCACTGATGGGATACGTGCATCGAGACATCAAGCTG 691
Db 663 TTGTCATGGCCATAGACTCGGTGACCGGCTTGCTACGTGCGAGGACATCAAAACCG 722
Qy 692 AGAATCTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTTGGATCTCGCCGGA 751
Db 723 ACAATCTCTGCTGACCGCTGTGGCCACATCCCGCTGGCGGACTTGGGCTCTTGGCTCA 782
Qy 752 AAATCAATTTCAACAAGATGTTGAATGCCAACTCCCGATTTGGGACCCCGAGATTACATGG 811
Db 783 AGCTGGGGCAGATGGAACGGTGGCTGCTGGTGGCTGGGACCCGACACTACCTGT 842
Qy 812 CTCCTGAAGTGTGA---CTGTGATGAACGGGGATGGAAAGGCACCTACGGCTGGACT 868

Db	843	CCCCGAGATCCCTGCAGGCTGTGGCGGTGGCGCTGGGACAGGCAGCTACGGGCCCAGT	902
Qy	869	GTGACTGGTGGTCACTGGGGCGTGATTCCTCATGAGATGATTATGGGAGATATCCCCCTTCG	928
Db	903	GTGACTGGTGGGCGCTGGGTGTATTGCCTATGAATGTTCTATGGCAGACGCCCTTCT	962
Qy	929	CAGAGGAACTCTGCCAGAACCTTCAATACATTATGAATTTCCAGCGGTTTTGAAAT	988
Db	963	ACGCGGATCCACGGCGGAGACCTATGGCAAGATCGCCACTCAAGAGGACACTCTCTC	1022
Qy	989	TTCCAGATGACCCCAAGTCAGGAGT---GACTTTCTTGATCTGATTCAAAGCTTTGTT	1045
Db	1023	TGCGGTGGTGACGAAGGGTCCCTGAGAGGCTCGAGACTTCATTCAGCGGTTGCTGT	1082
Qy	1046	CGCGCCAGAAGAGAGACTGAAGTTTGAAAGTCTTTGCTGC-----CATCTTTTCT	1096
Db	1083	GTCCCCGGGAGACACGGCTGGCGGGGTGGAGCAGGCGACTTCGAGACATATCCCTTCT	1142
Qy	1097	TCCTATAAATTGACTTGAACAACATTGTAACCTCTCTCCCGCCCTTCGTTCCACACCTCA	1156
Db	1143	TCATTGGCTCGACTGGGATGGTCTCCGGGACAGGTGCCCCCTTTACACCGGATTTTCG	1202
Qy	1157	AGTCTACGATGACACTCCAAATTTGA	1184
Db	1203	AAGTGCACCCGACATGTCAACTTGA	1230

RESULT 7

US-08-422-699A-8
Sequence 8, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422.699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:

QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAGCACT---GACTTCTTGATCTGATTCAA 1035
Db 763 CACCTCTCTCTCCGCTGTGTGACGAGGGGTCCCTGAGGAGGCTCGAGACTTTCATCAG 822
QY 1036 ACCTGTGTGTGGGGCCAGAGAGAGACTGAAGTTTGAAGGCTTTTGTCTG----- 1086
Db 823 CGTTGTGTGTCTCCCGGAGACAGGCTTGGCGGGGTGGAGCAGGCGACTTCGGGACA 882
QY 1087 CATCTTTCTTCTTAAATTTGACTGGAAACAATCTGTAATCTCTCTCCCTCCCTTCTGTT 1146
Db 883 CATCCCTTCTTCTTGGCTGCTGACTGGATGCTCTCCGGGACAGCGTCCCTCCCTTTACA 942
QY 1147 CCACCTCAAGTCTGACGATGACACCTCCCAATTTGA 1184
Db 943 CCGGATTTGGAAGTCCACCGACACATGCAACTTGA 980

RESULT 8

US-08-422-706B-8
Sequence 8, Application US/08422706B
Patent No. 5977333
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-706B-8
Query Match
Best Local Similarity 3.3%; Score 214; DB 2; Length 2511;
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;
QY 262 GACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAGTCAAGT 321
Db 43 GCCATCTGTGTAGGCTTAAGAGGTCGACTCGAGAGGACGACTTCGAGATTCGAAG 102
QY 322 CTTGTAGGTTGTGTCACTTTGTGAAGTGCAGGTGTTAAGAGAAAGCAACCGGGGAC 381
Db 103 GTGATCGGACCGGGGGGTTTCAAGGTTAGCTAGTGAAGATGAAGCAGAGCGGGCAG 162
QY 382 ATCTATCTATGAAGTATGATGAAGAGAGGCTTTATTGGCCAGGAGAGGTTTCAATT 441
Db 163 GTGTATGCCATGAAGATCATGAACAGTGGGACATGCTGAAGAGGGCGAGGTGCTGTC 222
QY 442 TTTGAGGAAGAGCGGAAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTAC 501
Db 223 TTCCGTGAGGAGGAGGACGTTTGGTGAATGGGACCGCGGTGGATCACGACGTCAC 282
QY 502 TATCCCTTTGAGGACAAATATCCTTTATCTGATGAGAGAAATATCAGCTTGGGGGAC 561
Db 283 TTCGCTTCCAGGATGAGAACTACCTGTACCTGTCTATGATGATGATGATGATGATG 342
QY 562 TTGCTGTCACTTTTGAATAGATATGAGGACCGAGTATGATGATGATGATGATGATGAT 621
Db 343 CTGCTGACATGCTGACCAAGTTTGGGAGGGGATTCGGCCGAGATGGCGGCTTCTAC 402
QY 622 CTAGCTGAGCTGATTTTGGCTGTTTCAAGCGGTTTCACTGATGAGGATGATGATGATG 681
Db 403 CTGGCGGAGATTGTCTGCGGATAGACTCGGTGTCACCGGCTTGGCTTGGCTACGTCAC 462
QY 682 ATCAAGCTGAGAACATCTCTGTCGCGCACAGGACACATCAAGCTGATGATGATGATG 741
Db 463 ATCAACCCGACAACTCTGCTGCGGACGCTGTGGCCACATCCGCTTGGCGGCTTGGC 522
QY 742 TCTGCGCGCAAAATGAATTTCAACCAAGATGATGATGATGATGATGATGATGATGATG 801
Db 523 TCTTGCCTCAAGCTCGGCGAGATGGAACGCTGCGTGGTGGCTGTGGGACACCCA 582
QY 802 GATTACATGGCTCTCTGA---AGTCTGACTGTGTGATGAACGGGGATGGAAGAGCCTAC 858
Db 583 GACTACCTGTCCCCCGAGATCCTGCGGCTGTGGGCGGTGGGCGCTGGGACAGGCGCTAC 642
QY 859 GGCCTGGACTGTGACTGTGTGCTGATGCTGCTGATGCTGCTGATGATGATGATGATGATG 918
Db 643 GGGCCGAGTGTGACTGGTGGCCCTGGGTGATTCGCTATGAAATGTTCTATGGGCGAG 702
QY 919 TCCCTCTTCGAGAGGGAACTCTGCGAGAACCTTCAATTAACATATGAATTTTCAGGG 978
Db 703 ACGCCCTTCTACGGCGATTCCACGGCGGAGACCTATGCGCAAGATCGTCCACTACAAGAG 762
QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGAGT---GACTTCTTGATCTGATTCAA 1035
Db 763 CACCTCTCTGCGGCTGTGTGGAAGAGGGTCCCTGAGGAGGCTCGAGACTTCATTCAG 822
QY 1036 AGCTTTGTTGTCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTTCTGCTG----- 1086
Db 823 CGGTGTGTGTCTCCCGGAGACACGCTGGGCGGGGTGGAGCAGGCGACTTCCGGACA 882
QY 1087 CATCTTTCTTCTTAAATTTGACTGGAAACAATCTGTAATCTCTCTCCCTCCCTTCTGTT 1146
Db 883 CATCCCTTCTTCTTGGCTGCTGACTGGATGCTCTCCGGGACAGCGTCCCTCCCTTTACA 942

QY 1147 CCCACCTCAAGTCTGACGATGACACCTCCAAATTTGA 1184
 Db 943 CCGGATTGCGAGGTGCCACCGACACATGCAACTTGA 980

RESULT 9

US-08-484-044-11
 ; Sequence 11, Application US/08484044
 ; Patent No. 5552282
 ; GENERAL INFORMATION:
 ; APPLICANT: Caskey, C. T.
 ; APPLICANT: Fu, Ying-Hui
 ; APPLICANT: Friedman, David L.
 ; APPLICANT: Pizzuti, Antonio
 ; APPLICANT: Penwick, Raymond G.
 ; TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fulbright & Jaworski, L.L.P.
 ; STREET: 1301 McKinney, Suite 5100
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: U.S.A.
 ; ZIP: 77010-3095
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,044
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/019,940
 ; FILING DATE: 19-FEB-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Paul, Thomas D.
 ; REGISTRATION NUMBER: 32,714
 ; REFERENCE/DOCKET NUMBER: D-5443
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 713/651-5325
 ; TELEFAX: 713/651-5246
 ; TELEX: 762829
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3182 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-484-044-11

Query Match 3.2%; Score 213.6; DB 1; Length 3182;
 Best Local Similarity 54.2%; Pred. No. 5.8e-47;
 Matches 507; Conservative 0; Mismatches 414; Indels 15; Gaps 3;

QY 264 CACATAGCTGAGTTACAGAGCTCCAGCTTCGGCAAGGACTTCGAAGTCAGAGTCT 323
 Db 743 CATCTGTGGTGGCTTAAGAGAGTCCGACTGCAGAGGGACGACTTCGAGATTCTGAAGGT 802
 QY 324 TGTAGTTGTGTCATTTGCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGGGACAT 383
 Db 803 GATCGACCGGGGCTTCAGCGAGGTAGCGGTAGTGAAGTGAAGCAGCGGGCCAGGT 862
 QY 384 CTATGCTATCAAGTCATGAAGAGCGCTTATGGCCAGGAGCGGTTTCAATTTT 443
 Db 863 GTATGCTATGAAGTCATGAAGAGCGCTTATGGCCAGGAGCGGTTTCAATTTT 922
 QY 444 TGAGGAGAGCGGAAACATATTATCTGAAGCACAAGCGCGTGGATCCCGCCCAATACAGTA 503
 Db 923 CCGTGAGGAGAGGAGCGGTGTTGTTGATCGGACCGCGGTGGATCAGCGACCTGCATT 982

QY 504 TGCCTTTCCAGGACAAAATCACCTTTATCTGATGAGGAATATCAGCTTGAGGGGACTT 563
 Db 983 CCGCTTTCCAGGATGAGAACTACCTGTACCTGGTTCATGGAGTATTACGTGGGCGGGGACCT 1042
 QY 564 GCTGTCACTTTTGAATAGATATGAGGACCAAGTATAGATGAAAACCTGTATACAGTTTACCT 623
 Db 1043 GCTGACACTGCTGACCAAGTTTGGGAGCGGATCCGCGCCGAGATGGCGCGCTTCTACT 1102
 QY 624 AGCTGAGCTGATTTGGCTGTTTCACAGCGTTTCATCTGATGGGATACGTCGATTCGAGACAT 683
 Db 1103 GCGCGAGATTGTCATGGCCATAGACTCGGTGCACCGGCTTGGCTACGTCGACAGGGACAT 1162
 QY 684 CAAGCTCGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGGATTTTGGATC 743
 Db 1163 CAAGCCCGACACATCTGCTGGACCGCTGTGGCCACATCCGCTGGCCGACTTGGGCTC 1222
 QY 744 TGCCCGGAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACCCGAGA 803
 Db 1223 TTGCCTCAAGCTGCGGGCAGATGGAACGGTCCGCTCGCTGTGGCGACCCGAGA 1282
 QY 804 TTACATGCTCCTGA---AGTCTGACTGTGATGAACGGGATGGAAGGACGACCTACGG 860
 Db 1283 CTACTGTCCCGGAGATCCTGCAGGCTGTGGCGGTGGGCTGGGACAGCAGCTACGG 1342
 QY 861 CTTGACTGTGACTGGTGGTCAAGTGGGCGTGAATTCCTATGAGATGATTTATGGGAGATC 920
 Db 1343 GCGGAGTGTGACTGGTGGGCGTGGTGTATTTCGCTATGAATGTTCTATGGGACAGC 1402
 QY 921 CCCCTTCGACAGGGAACCTCTGCAGAACCTTCAATACATATGAATTTCCAGCGGTT 980
 Db 1403 GCGCTTCTACCGGATTCACGCGCGGAGACTTATGGCAAGATCGTCCACTACAGGAGCA 1462
 QY 981 TTTGAAATTTCCAGATGACCCCAAAGTGAGCAGT---GACTTTCTTGTATCTGATTTCAAAG 1037
 Db 1463 CCTCTCTCGCGCTGGTGGAGGAGGGGTCCCTGAGGAGGCTCGAGACTTCATTACGG 1522
 QY 1038 CTTGTTGGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGTGTC-----CA 1088
 Db 1523 GTTGTGTGTCCCGGAGACACGCGTGGCGCGGGTGGAGCAGGCGACTTCGCGGACACA 1582
 QY 1089 TCCTTTCTTCTCTAAATTCAGTGGAAACAACATTCGPAACTCTCTCCCTCCCGCTTCGTTCC 1148
 Db 1583 TCCCTTTCTTCTTGGCGCTGACTGGGATGGTCTCCGGGACAGCGTGTCCCTTTACACC 1642
 QY 1149 CACCTCAAGTCTGACGATGACACCTCCAAATTTGA 1184
 Db 1643 GGATTCGAAGGTGCCACCGACACATGCAACTTCGA 1678

RESULT 10

US-09-804-471A-3
 ; Sequence 3, Application US/09804471A
 ; Patent No. 6479269
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marion et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001164
 ; CURRENT APPLICATION NUMBER: US/09/804,471A
 ; CURRENT FILING DATE: 2001-03-13
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 174493
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(174493)
 ; OTHER INFORMATION: n = A, T, C or G
 ; US-09-804-471A-3

```
Query Match
Best Local Similarity 3.1%; Score 205; DB 4; Length 174493;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 771 GGTGAATGCCAAACTCCCGATTGGGACCCCGACATTACATGGCTCCTGAAGTCTGACTGT 830
Db 130289 GGTGAATGCCAAACTCCCGATTGGGACCCCGACATTACATGGCTCCTGAAGTCTGACTGT 130348

QY 831 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGTGGTGGTGGCGGT 890
Db 130349 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGTGGTGGTGGCGGT 130408

QY 891 GATTCCCTATGAGATGATTATGGGAGATCCCTCCAGAGGAACTCTGCCAGAAC 950
Db 130409 GATTCCCTATGAGATGATTATGGGAGATCCCTCCAGAGGAACTCTGCCAGAAC 130468

QY 951 CTTCAATAACATTATGAATTTCCAG 975
Db 130469 CTTCAATAACATTATGAATTTCCAG 130493

RESULT 11
US-10-238-709-3
; Sequence 3, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Query Match
Best Local Similarity 3.1%; Score 205; DB 4; Length 174493;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 771 GGTGAATGCCAAACTCCCGATTGGGACCCCGACATTACATGGCTCCTGAAGTCTGACTGT 830
Db 130289 GGTGAATGCCAAACTCCCGATTGGGACCCCGACATTACATGGCTCCTGAAGTCTGACTGT 130348

QY 831 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGTGGTGGTGGCGGT 890
Db 130349 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGTGGTGGTGGCGGT 130408

QY 891 GATTGCTATGAGATGATTATGGGAGATCCCTCCAGAGGAACTCTGCCAGAAC 950
Db 130409 GATTGCTATGAGATGATTATGGGAGATCCCTCCAGAGGAACTCTGCCAGAAC 130468

QY 951 CTTCAATAACATTATGAATTTCCAG 975
Db 130469 CTTCAATAACATTATGAATTTCCAG 130493

RESULT 12
US-08-630-822A-61
; Sequence 61, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
```

```
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDIA
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
City: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,822A
FILING DATE: 11-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 2706 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 5..2706
US-08-630-822A-61

Query Match
Best Local Similarity 2.9%; Score 189.2; DB 2; Length 2706;
Matches 489; Conservative 0; Mismatches 398; Indels 27; Gaps 4;

QY 299 CAAAGGACTTCGAAGTCAGAAAGTCTTGTAGTGTGGTCACTTTGCTGGAAGTGCAGGTGG 358
Db 78 CAGATGATTTTAAATTTAATAAAGTTATTGCTCGAGGAGCATTTGTTGAAAGTACAGTTAG 137

QY 359 TAAGAGAGAACCAACCGGGGACATCTATGCTATGAAGTGATGAAGAGAGGCTTTAT 418
Db 138 TGGACACAAATCAACTGCACAAAGTTTTCGTATGAACCCCTATCAAAATTTGAAATGA 197

QY 419 TGGCCAGGAGCAGGTTTCAATTTTTCAGGAAGAGCGGAAACATATTATCTCGAAGCACA 478
Db 198 TTAAGAGACCAGACTCTGCATTTTTCAGGAAGACGTCATATAATGGCTCATGCAAAAT 257

QY 479 GCCCGTGATCCCAATTCAGTATGCCCTTTCAGGCAAAATACCTTTATCTGATGG 538
Db 258 CAGAATCGATTGTACAATTCATTTTTCGCTTTCAGAGATCAAAATATCTTTATATGTC 317

QY 539 AGGAATATCAGCTCGAGGGGACTTGTCTGTCACCTTTTGAATAGATATGAGCAAGTTAG 598
Db 318 TGGATTATATGCCGGGGGTGACTTGGTGAATCTTATG-----TCCGATTATGAATTC 371

QY 599 ATGAAACCTGATACAGTTTACCTAGCTGAGTGTATTTGGCTGTTTCAGCGTTCATC 658
Db 372 CAGAAAAATGGCAATGTTCTATACAAATGGAAGTGGTGTAGCACTTGATACAAATTCAC 431

QY 659 TGATGGGATACGTCATCGAGACATCAAGCTCAGAACATTCCTGTTGACCGCACAGAC 718
Db 432 CCATGGGATTTGACATCGTGATGTTAACTGTTAACTGATTAATGCTTCTAGACAAATATG 491

QY 719 ACATCAAGCTGGTGGATTTTGGATCTCGCGGAAATGAATTCAAACAGATGGTGAAT 778
```


Search completed: July 4, 2004, 09:51:43
Job time : 402 secs

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 18:44:52 ; Search time 2174 Seconds
(without alignments)
12846.207 Million cell updates/sec

Title: US-10-017-216-1
Perfect score: 6574
Sequence: 1 agagccgcagtgaggagat.....atccagaatgaggtttaga 6574

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_29Jan04.*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002as.*
 - 7: Geneseqn2003as.*
 - 8: Geneseqn2003bs.*
 - 9: Geneseqn2003cs.*
 - 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6574	100.0	6574	6 AAD39191	Aad39191 Human MDP
2	5684.4	86.5	6298	6 AAD38864	Aad38864 Human kin
3	5668	86.2	8603	8 AAL55215	Aal55215 Human CRI
4	5666.6	86.2	6156	8 AAL55217	Aal55217 Human CRI
5	5666.2	86.2	6165	8 AAL55214	Aal55214 Human CRI
6	5661.4	86.1	6165	6 ABQ78870	Abq78870 Human kin
7	5650.4	86.0	6159	6 AAS06701	Aas06701 Polynucle
8	5631	85.7	6189	6 ABS63436	Abs63436 RHO/RAC-1
9	5631	85.7	6189	7 ADA05641	Ada05641 Human NOV
10	5629.8	85.6	6201	6 ABS63435	Abs63435 Human cDN
11	5629.8	85.6	6201	7 ADA05653	Ada05653 Human NOV
12	5358	81.5	5877	6 ABQ78871	Abq78871 Human kin
13	3475.2	52.9	6809	3 AAC77568	Aac77568 Human ORF
14	2575.2	39.2	2896	5 ABV30132	Abv30132 Human pro
15	2564.8	39.0	3131	4 ABA08361	Aba08361 Human RHO
16	2436.8	37.1	5261	8 AAL55216	Aal55216 Human CRI
17	2426.8	36.9	5251	9 ADD89966	Add89966 Human can
18	2414	36.7	2542	7 ADA05647	Ada05647 Human NOV
19	2317.2	35.2	2497	7 ADA05645	Ada05645 Human NOV
20	2126.8	32.4	2893	9 ADS09823	Ads09823 Novel DNA
21	1841.2	28.0	1870	7 ADA05643	Ada05643 Human NOV
22	1524.2	23.2	1870	7 ADA05649	Ada05649 Human NOV
23	1467.6	22.3	1915	7 ADA05651	Ada05651 Human NOV

ALIGNMENTS

RESULT 1

AAD39191 ID AAD39191 standard; cDNA; 6574 BP.

XX AC AAD39191;

XX DT 04-OCT-2002 (first entry)

XX DE Human MDPK cDNA.

XX KW Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;
tumorigenesis; tumour growth; tumour metastasis; viral infection;
skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;
immune disorder; neoplastic disorder; gene therapy; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT 5'UTR 1..18

FT CDS /*tag= a

FT /*tag= b

FT misc_feature /product= "Human MDPK protein"

FT /*tag= c

FT /*tag= d

FT 3'UTR as SEQ ID NO:3"

XX WO200234896-A2.

XX 02-MAY-2002.

XX 23-OCT-2001; 2001WO-US050636.

XX 23-OCT-2000; 2000US-0242429P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kapeller-Libermann R;

XX WPI; 2002-479720/51.

XX P-PSDB; AAE24079.

XX Human myotonic dystrophy type protein kinase polypeptide and

PT polynucleotide useful for prognosticating, diagnosing, preventing or
PT inhibiting tumorigenesis, tumor growth, tumor metastasis and viral
XX infection.

PS Claim 1; Fig 1; 148pp; English.

XX The invention relates to human myotonic dystrophy type protein kinase
CC (MDPK) polypeptides designated as 13245 and nucleic acid molecules
CC encoding such polypeptides. 13245 molecules are used to develop
CC diagnostic and therapeutic agents for prognosticating, diagnosing,
CC preventing, inhibiting, alleviating or curing MDPK-related disorders.
CC Polypeptides of the invention are used to develop diagnostic and
CC therapeutic agents for 13245-mediated or related disorders such as
CC tumorigenesis, tumour growth, tumour metastasis, viral infection of a
CC cell, skeletal muscle disorders (e.g. muscular and myotonic dystrophies),
CC immune disorders and neoplastic disorders. The invention is also used in
CC gene therapy. The present sequence is human MDPK cDNA

XX Sequence 6574 Bp; 1877 A; 1611 C; 1776 G; 1310 T; 0 U; 0 Other;

Query Match	100.0%;	Score 6574;	DB 6;	Length 6574;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6574;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGGAGCGGGAATCCTTTGGATGCT	60		
DB 1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGGAGCGGGAATCCTTTGGATGCT	60		
QY 61	GGTGTCTGTAACCCATTTGCGACCGCGGCTCAGGCTGAATCTGTTCTCCAGGGAAA	120		
DB 61	GGTGTCTGTAACCCATTTGCGACCGCGGCTCAGGCTGAATCTGTTCTCCAGGGAAA	120		
QY 121	CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATGC	180		
DB 121	CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATGC	180		
QY 181	CTCTTTGTTCTTTGAAGATCAGTCAGCCTGCTCTGATGAAGATTAGCAGCTGAGC	240		
DB 181	CTCTTTGTTCTTTGAAGATCAGTCAGCCTGCTCTGATGAAGATTAGCAGCTGAGC	240		
QY 241	AATTTGTCGGAGATTTCGGACACATAGCTAGTACAGAGCTCCAGCCTTCGGCA	300		
DB 241	AATTTGTCGGAGATTTCGGACACATAGCTAGTACAGAGCTCCAGCCTTCGGCA	300		
QY 301	AAGGACTTCGAAGTCAAGATCTTTAGGTTGTTGTTGCTTCTGCTGAAGTCAAGTGGTA	360		
DB 301	AAGGACTTCGAAGTCAAGATCTTTAGGTTGTTGTTGCTTCTGCTGAAGTCAAGTGGTA	360		
QY 361	AGAGAGAACCAACCGGGACATCTATGTCATGAAGTATGAGNAGAGGCTTTATTG	420		
DB 361	AGAGAGAACCAACCGGGACATCTATGTCATGAAGTATGAGNAGAGGCTTTATTG	420		
QY 421	GCCAGGAGCAGGTTTCATTTTGAAGAGAGCGGAACATATTATCTCAAGCAACAGC	480		
DB 421	GCCAGGAGCAGGTTTCATTTTGAAGAGAGCGGAACATATTATCTCAAGCAACAGC	480		
QY 481	CCGTGATCCCCCAATTACAGTATGCTTTTCAAGCAAAATATCCTTTATCTGATGGAG	540		
DB 481	CCGTGATCCCCCAATTACAGTATGCTTTTCAAGCAAAATATCCTTTATCTGATGGAG	540		
QY 541	GAATATCAGCTCGAGGGGACTTGTCTGCTCACTTTTGAATAGATATGAGGACAGTTAGAT	600		
DB 541	GAATATCAGCTCGAGGGGACTTGTCTGCTCACTTTTGAATAGATATGAGGACAGTTAGAT	600		
QY 601	GAATACTGATACAGTTTACCTAGCTGAGTGAATTTGGCTGTTTCAAGCTTCACTG	660		
DB 601	GAATACTGATACAGTTTACCTAGCTGAGTGAATTTGGCTGTTTCAAGCTTCACTG	660		
QY 661	ATGGGATACGTGATCGAGACATCAAGCTGAGACATCTCGTTCAGCCGACAGACAC	720		
DB 661	ATGGGATACGTGATCGAGACATCAAGCTGAGACATCTCGTTCAGCCGACAGACAC	720		
QY 721	ATCAAGCTGGTGGATTTTGGATCTGCCGGAATTAATTCAAACAGATGTTGAATGCC	780		

DB 721	ATCAAGCTGGTGGATTTTGGATCTGCCGGAATTAATTCAAACAGATGTTGAATGCC	780
QY 781	AAACTCCCGATTGGACCCCCAGATTACATGGCTCCTGAAGTGTGATGTGAACGGG	840
DB 781	AAACTCCCGATTGGACCCCCAGATTACATGGCTCCTGAAGTGTGATGTGAACGGG	840
QY 841	GATGAAAGCCACTACGCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900
DB 841	GATGAAAGCCACTACGCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900
QY 901	GAGATGATTTATGGAGATCCCTCTTCGACAGAGGAACCTCTGCCAGAACCTTCAATAAC	960
DB 901	GAGATGATTTATGGAGATCCCTCTTCGACAGAGGAACCTCTGCCAGAACCTTCAATAAC	960
QY 961	ATTATGATTTCCAGCGGTTTGTGAATTTCCAGATGACCCCAAGTGGACGACTGCTT	1020
DB 961	ATTATGATTTCCAGCGGTTTGTGAATTTCCAGATGACCCCAAGTGGACGACTGCTT	1020
QY 1021	CTTGATCTGATTTCAAAGCTTGTGCGGCGCAGAAAGAGAGACTGAAAGTTTGAAGTCTT	1080
DB 1021	CTTGATCTGATTTCAAAGCTTGTGCGGCGCAGAAAGAGAGACTGAAAGTTTGAAGTCTT	1080
QY 1081	TGCTGCCATCTTTCTCTCTAAAATTGACTGGAACAACTTCTCTCTCTCTCTCTCT	1140
DB 1081	TGCTGCCATCTTTCTCTCTAAAATTGACTGGAACAACTTCTCTCTCTCTCTCTCT	1140
QY 1141	TTCTGTTCCACCTTCAAGTCTGACGATGACCTTCAATTTTGTATGAACCAAGAGAAAT	1200
DB 1141	TTCTGTTCCACCTTCAAGTCTGACGATGACCTTCAATTTTGTATGAACCAAGAGAAAT	1200
QY 1201	TCGTGGGTTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
DB 1201	TCGTGGGTTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
QY 1261	CCGTGTTGGGGTTTCTGACAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
DB 1261	CCGTGTTGGGGTTTCTGACAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
QY 1321	GTGTGGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1380
DB 1321	GTGTGGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1380
QY 1381	AGCAAGAGCTTACAGACTCTCTGAGCAAGTGTCTCAAGATGGAGAGGAAATGACCCG	1440
DB 1381	AGCAAGAGCTTACAGACTCTCTGAGCAAGTGTCTCAAGATGGAGAGGAAATGACCCG	1440
QY 1441	TTACATCGGAGTGTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
DB 1441	TTACATCGGAGTGTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
QY 1501	GCCTCTGAGACTCAGAGATCCCTCTCTGAGCAGGACTTCTGCTACATCAAGATGTC	1560
DB 1501	GCCTCTGAGACTCAGAGATCCCTCTCTGAGCAGGACTTCTGCTACATCAAGATGTC	1560
QY 1561	AGTACGTTAAAGCGAAGTTTGGAGCAAGCAGGATGGAGTGTCTGCTGCTGCTGCTG	1620
DB 1561	AGTACGTTAAAGCGAAGTTTGGAGCAAGCAGGATGGAGTGTCTGCTGCTGCTGCTG	1620
QY 1621	GCACTGAGCTTCTCCATGATATCAGAGACAGAGCCGGAAGCTTCAAGAAATCAAGAG	1680
DB 1621	GCACTGAGCTTCTCCATGATATCAGAGACAGAGCCGGAAGCTTCAAGAAATCAAGAG	1680
QY 1681	CAGAGTACCGGCTCAAGTGGAGAAATGAGGTTGATGATGATGATGATGATGATGATG	1740
DB 1681	CAGAGTACCGGCTCAAGTGGAGAAATGAGGTTGATGATGATGATGATGATGATGATG	1740
QY 1741	CTTCTCTCAGCAAGAGACGAGTGTCTTACGAATCTCAGCTGAGAGAGTCTCGGCTT	1800
DB 1741	CTTCTCTCAGCAAGAGACGAGTGTCTTACGAATCTCAGCTGAGAGAGTCTCGGCTT	1800
QY 1801	GCTCTGAGGATTCAGCGGAAGCCAGCAATGTCAGCATTAATCTGTTGAGGCTTAAG	1860

Db 1801 GCTGCTGAAGAAATTCAGCGGAAGCGACAGAAATGTACGATATAAACTGTTGAAGCTAAG 1860
Qy 1861 GATCAAGGGAAGCCTGAAGTGGGAGAAATATGCGAAACTGAGAGAAATCAATGCTCAGCAG 1920
Db 1861 GATCAAGGGAAGCCTGAAGTGGGAGAAATATGCGAAACTGAGAGAAATCAATGCTCAGCAG 1920
Qy 1921 CAGCTCAAAATTCAGGAGTCCAAAGAGAACTGGAGAGGCTGCAAGAGCGAGCCGAG 1980
Db 1921 CAGCTCAAAATTCAGGAGTCCAAAGAGAACTGGAGAGGCTGCAAGAGCGAGCCGAG 1980
Qy 1981 AGGAGCTCGAGAGAGCTGAGAGACGAGAGGATCTCTGAGAGGATCAGAAAGAGCTG 2040
Db 1981 AGGAGCTCGAGAGAGCTGAGAGACGAGAGGATCTCTGAGAGGATCAGAAAGAGCTG 2040
Qy 2041 GTGGAAGCTGAGGAACCGCGCCATCTCTGGAAGCAAGAGTAAAGAGACTAGAGACCATG 2100
Db 2041 GTGGAAGCTGAGGAACCGCGCCATCTCTGGAAGCAAGAGTAAAGAGACTAGAGACCATG 2100
Qy 2101 GAGCTAGAGAAAACAGACTGAGGATGATATCCAGAGAAAATCCCAAGATCCAGAG 2160
Db 2101 GAGCTAGAGAAAACAGACTGAGGATGATATCCAGAGAAAATCCCAAGATCCAGAG 2160
Qy 2161 ATGGCTGATAAAATCTGGAGTTCGAGAGAAACATCGGAGGCGCCCAAGTCTCAGCCAG 2220
Db 2161 ATGGCTGATAAAATCTGGAGTTCGAGAGAAACATCGGAGGCGCCCAAGTCTCAGCCAG 2220
Qy 2221 CACCTAGAGTGCACCTGAAAACAGAAAGAGAGCACTATGAGGAAAAGATTAAAGTGTG 2280
Db 2221 CACCTAGAGTGCACCTGAAAACAGAAAGAGAGCACTATGAGGAAAAGATTAAAGTGTG 2280
Qy 2281 GACATACATATAAGAAAGACCTGGCTGACAAAGGAGACATCGAGAAACATGATGAGAGA 2340
Db 2281 GACATACATATAAGAAAGACCTGGCTGACAAAGGAGACATCGAGAAACATGATGAGAGA 2340
Qy 2341 CACGAGGAGGCGCCATGAGAGGCAAAAATCTCAGCGAAACAGAAAGCGATGATCAAT 2400
Db 2341 CACGAGGAGGCGCCATGAGAGGCAAAAATCTCAGCGAAACAGAAAGCGATGATCAAT 2400
Qy 2401 GCTATGATTCAGAGTACAGATCCCTGGAGAGAGATGAGAGATGAGTGTGAGCCCAAT 2460
Db 2401 GCTATGATTCAGAGTACAGATCCCTGGAGAGAGATGAGAGATGAGTGTGAGCCCAAT 2460
Qy 2461 AAACCTGACGAAATAGCAGTCTTTTACCAGAGAAACATGAAGGCGCCCAAGAGAGATG 2520
Db 2461 AAACCTGACGAAATAGCAGTCTTTTACCAGAGAAACATGAAGGCGCCCAAGAGAGATG 2520
Qy 2521 ATTTCTGAATCAGGCAACAGAAAATTTTACCTGGAGACACAGGCTGGAGTTGAGGCC 2580
Db 2521 ATTTCTGAATCAGGCAACAGAAAATTTTACCTGGAGACACAGGCTGGAGTTGAGGCC 2580
Qy 2581 CAGAACCGAAAACCTGGAGGAGCAGCTGGAGAGATCAGCCACAGAGCCACAGTGACAAAG 2640
Db 2581 CAGAACCGAAAACCTGGAGGAGCAGCTGGAGAGATCAGCCACAGAGCCACAGTGACAAAG 2640
Qy 2641 AATCGGCTGCTGGAATCGAGACAAAGATTGCGGAGGTCAGTCTAGACACAGAGAGCAG 2700
Db 2641 AATCGGCTGCTGGAATCGAGACAAAGATTGCGGAGGTCAGTCTAGACACAGAGAGCAG 2700
Qy 2701 AACTGGAGCTCAAGGCCAGCTCAGAGCTACAGCTCTCCTCGAGAGCGGAGTCA 2760
Db 2701 AACTGGAGCTCAAGGCCAGCTCAGAGCTACAGCTCTCCTCGAGAGCGGAGTCA 2760
Qy 2761 CAGTTGACAGCCCTGAGGCTGCAAGGCGGCGCTGGAGAGCAGGCTTCGCGAGCGGAAG 2820
Db 2761 CAGTTGACAGCCCTGAGGCTGCAAGGCGGCGCTGGAGAGCAGGCTTCGCGAGCGGAAG 2820
Qy 2821 ACAGAGCTGGAAGAGACACAGAGAGCTGAAGAGAGATCCAGGCACTCAGGCAAT 2880
Db 2821 ACAGAGCTGGAAGAGACACAGAGAGCTGAAGAGAGATCCAGGCACTCAGGCAAT 2880
Qy 2881 AGAGATGAATCCAGGCGAAAATTTGATGCTCTCTGTAACAGCTGTACTGTAAATCAGAC 2940
Db 2881 AGAGATGAATCCAGGCGAAAATTTGATGCTCTCTGTAACAGCTGTACTGTAAATCAGAC 2940

Qy 2941 CTGAGGAGCAGCTAAACCAGCTGACCGAGGACAAACGCTGAACCTCAACAAACCAAACTTC 3000
Db 2941 CTGAGGAGCAGCTAAACCAGCTGACCGAGGACAAACGCTGAACCTCAACAAACCAAACTTC 3000
Qy 3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGGCCCAACGACGAGATTGTACAACTCGA 3060
Db 3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGGCCCAACGACGAGATTGTACAACTCGA 3060
Qy 3061 AGTGAAGTGAACCATCTCCGCGGAGATCAAGAAACGAGAGATGACGCTTACCAGCAG 3120
Db 3061 AGTGAAGTGAACCATCTCCGCGGAGATCAAGAAACGAGAGATGACGCTTACCAGCAG 3120
Qy 3121 AAGCAAAACGATGGAGGCTCTGAAGACCAACGTCACCACTGTGGAAGGAACAGGTCAATGAT 3180
Db 3121 AAGCAAAACGATGGAGGCTCTGAAGACCAACGTCACCACTGTGGAAGGAACAGGTCAATGAT 3180
Qy 3181 TTGAGAGCCCTAAACGATGAGCTCTAGAAAAGAGCGGAGTGGAGGCTTGGAGAGC 3240
Db 3181 TTGAGAGCCCTAAACGATGAGCTCTAGAAAAGAGCGGAGTGGAGGCTTGGAGAGC 3240
Qy 3241 GTCTGTGATGATGAGAAATCCAGTTTGAAGTGTGCGGTTGAGAGCTGCGAGAGATGCTG 3300
Db 3241 GTCTGTGATGATGAGAAATCCAGTTTGAAGTGTGCGGTTGAGAGCTGCGAGAGATGCTG 3300
Qy 3301 GACACCGAGAAAACAGAGAGGCGAGAGCCGATCAGCGATCACCGAGTCTCGCAGGTG 3360
Db 3301 GACACCGAGAAAACAGAGAGGCGAGAGCCGATCAGCGATCACCGAGTCTCGCAGGTG 3360
Qy 3361 GTGAGCTGGCAGTGAAGGAGCACAAGGCTGAGATTCTCGCTCTGCGAGAGGCTCTCAAA 3420
Db 3361 GTGAGCTGGCAGTGAAGGAGCACAAGGCTGAGATTCTCGCTCTGCGAGAGGCTCTCAAA 3420
Qy 3421 GAGCAGAAGCTGAAGCGGAGAGCCTCTCTGACAAGCTCAATGACCTGGGAGAAAGCAT 3480
Db 3421 GAGCAGAAGCTGAAGCGGAGAGCCTCTCTGACAAGCTCAATGACCTGGGAGAAAGCAT 3480
Qy 3481 GCTATGCTTGAATGAATGCTTACAGAGAGCTTACAGAGAGCTGGAAGCTGGAAGAGCTC 3540
Db 3481 GCTATGCTTGAATGAATGCTTACAGAGAGCTTACAGAGAGCTGGAAGCTGGAAGAGCTC 3540
Qy 3541 AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTTACAGAGAGATGGAACCTGCGAGAAAAT 3600
Db 3541 AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTTACAGAGAGATGGAACCTGCGAGAAAAT 3600
Qy 3601 CACATTTTCGCTGATCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTACTGAAG 3660
Db 3601 CACATTTTCGCTGATCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTACTGAAG 3660
Qy 3661 ACAGAAAGAGTGAATTTGAGATTCAGCTGGAAGCAATTCAGGTTCTCTATTCTCATGAA 3720
Db 3661 ACAGAAAGAGTGAATTTGAGATTCAGCTGGAAGCAATTCAGGTTCTCTATTCTCATGAA 3720
Qy 3721 AAGTGAAAATGGAAGCAGCTATTTCTCAACAAACCAAACTCATTTCTGCAAGCC 3780
Db 3721 AAGTGAAAATGGAAGCAGCTATTTCTCAACAAACCAAACTCATTTCTGCAAGCC 3780
Qy 3781 AAATGGAACCACTGCTTAAAGAAAAGGGTTATTTAGTTCGACGGAAGAGGACCT 3840
Db 3781 AAATGGAACCACTGCTTAAAGAAAAGGGTTATTTAGTTCGACGGAAGAGGACCT 3840
Qy 3841 GCTTTACCCACAGGTTCTCTGCAATCAATGAGCTGAAGCTGGCCCTTGGAGAGGAG 3900
Db 3841 GCTTTACCCACAGGTTCTCTGCAATCAATGAGCTGAAGCTGGCCCTTGGAGAGGAG 3900
Qy 3901 AAAGCTGCTGTCAGAGCTAGAGGAGCCCTTCAAGAACCCGATCAGCTCCCGTCC 3960
Db 3901 AAAGCTGCTGTCAGAGCTAGAGGAGCCCTTCAAGAACCCGATCAGCTCCCGTCC 3960
Qy 3961 GCCCGGAGGAGCTGCCACCGCAAGCAACCGAGCAACCCACACCCATCCAGCCAGCC 4020
Db 3961 GCCCGGAGGAGCTGCCACCGCAAGCAACCGAGCAACCCACACCCATCCAGCCAGCC 4020

Db 6181 GCAGAGTTCATGCTCTTACAGCTGGTGACCTTAAAAATGCGCTTAAAGCTGCAGAGC 6240
QY 6241 CAGCCACCTCTCTTACAAAAAGAGTACTTACTGTCACATGCTGTAGCAAAACAATTGTAA 6300
Db 6241 CAGCCACCTCTCTTACAAAAAGAGTACTTACTGTCACATGCTGTAGCAAAACAATTGTAA 6300
QY 6301 AACCTCATCTGAATACAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360
Db 6301 AACCTCATCTGAATACAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360
QY 6361 GAGAGACAATCTGTTGTTGATTTTGAAGACAGGCAAGCAACACTCTATTAGTTTCCA 6420
Db 6361 GAGAGACAATCTGTTGTTGATTTTGAAGACAGGCAAGCAACACTCTATTAGTTTCCA 6420
QY 6421 TAGCCAGGCTCAACAGGACAGTGGCTGGGCTTTAAAAACACACAGATGACTGGAATG 6480
Db 6421 TAGCCAGGCTCAACAGGACAGTGGCTGGGCTTTAAAAACACACAGATGACTGGAATG 6480
QY 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGAGGATAGCATTCATT 6540
Db 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGAGGATAGCATTCATT 6540
QY 6541 TTGGCTTAAGAAAAATCGAAGATGATGTTTAGA 6574
Db 6541 TTGGCTTAAGAAAAATCGAAGATGATGTTTAGA 6574

RESULT 2

AAD38864
ID AAD38864 standard; cDNA; 6298 BP.

XX AAD38864;

XX 23-SEP-2002 (first entry)

XX Human kinase (PKIN)-21 cDNA.

XX Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
KW development; hepatitis; cardiovascular; hypertension; drug screening;
KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
KW hypercholesterolemia; obesity; gene therapy; cytostatic; anti-HIV;
KW neuroprotective; hepatotropic; hypotensive; cardiac; nephrotropic;
KW hyperlipidaemia; enzyme; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 55..6219
CDS /*tag= a

FT /product= "Human kinase (PKIN)-21"

XX WC200233099-A2.

XX 25-APR-2002.

XX 20-OCT-2001; 2001WC-US047728.

XX 20-OCT-2000; 2000US-0242410P.

XX 27-OCT-2000; 2000US-0244068P.

XX 03-NOV-2000; 2000US-0245708P.

XX 09-NOV-2000; 2000US-0247672P.

XX 16-NOV-2000; 2000US-0249565P.

XX 22-NOV-2000; 2000US-0252730P.

XX 01-DEC-2000; 2000US-0250807P.

XX (INCY-) INCVTE GENOMICS INC.

XX Gururajan R, Baughn MR, Walia NK, Elliott VS, Xu Y, Arvizu C;

PI Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;

PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;

PI Recipon SA, Lu DAM, Borowsky ML, Thornton M, Swarnaker A;
PI Thangavelu K, Khan FA, Ison CH;
XX WPI; 2002-454603/48.
DR P-PSDB; AAE24150.

XX New human kinase polypeptide, for diagnosing, preventing and treating
PT cancer, immune system disorders, growth and development disorders,
PT cardiovascular disorders and lipid disorders.

XX Claim 5; Page 207-209; 210pp; English.

XX The invention relates human kinases (PKIN) and their corresponding
CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
CC treating and preventing cancer, an immune system disorder (e.g., acquired
CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
CC cardiovascular disorder (e.g., hypertension, myocardial infarction,
CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,
CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
CC condition or a disease associated with the expression of PKIN in a
CC biological sample. A composition comprising PKIN or an agonist or
CC antagonist of PKIN is useful for treating a disease or condition
CC associated with decreased or increased expression of functional PKIN.
CC PKIN is useful in a number of drug screening techniques and to analyse
CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
CC knockin humanised animals or transgenic animals to model human diseases,
CC and in somatic or germline gene therapy. The present sequence is human
CC PKIN cDNA

XX Sequence 6298 BP; 1772 A; 1585 C; 1720 G; 1221 T; 0 U; 0 Other;

Query Match 86.5%; Score 5684.4; DB 6; Length 6298;

Best Local Similarity 95.5%; Pred. No. 0;

Matches 5979; Conservative 0; Mismatches 36; Indels 243; Gaps 3;

QY 1 AGAGCCGCCAGTGGGAGATGTTGAAGTTCAAAATATGGAGCGGGAATCTTTGGATGCT 60

Db 37 AGAGCCGCCAGTGGGAGATGTTGAAGTTCAAAATATGGAGCGGGAATCTTTGGATGCT 96

QY 61 GGTGCTCTGAAACCCATATGTCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAA 120

Db 97 GGTGCTCTGAAACCCATATGTCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAA 156

QY 121 CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAAGGATATTAGATGCC 180

Db 157 CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAAGGATATTAGATGCC 216

QY 181 CTCTTTCTTCTCTTTGAAGAATGACGTACGCTCTCTGATGAAGATTAAGCAGTGAGC 240

Db 217 CTCTTTCTTCTTTGAAGAATGACGTACGCTCTCTGATGAAGATTAAGCAGTGAGC 276

QY 241 AACTTTGTCGGAAGTATTCGACACCATAGCTAGGTACAGGAGCTCCAGCTTCGGCA 300

Db 277 AACTTTGTCGGAAGTATTCGACACCATAGCTAGGTACAGGAGCTCCAGCTTCGGCA 336

QY 301 AAGGACTTCGAAGTCAGAAGTCTTTAGGTTGGTGCACCTTGTGCTGAAGTCAGGTGCTA 360

Db 337 AAGGACTTCGAAGTCAGAAGTCTTTAGGTTGGTGCACCTTGTGCTGAAGTCAGGTGCTA 396

QY 361 AGAGAGAAACAAACCGGGGACATCTATGCTATGAAGTGTATGAAGAAGAGGCTTTATTG 420

Db 397 AGAGAGAAACAAACCGGGGACATCTATGCTATGAAGTGTATGAAGAAGAGGCTTTATTG 456

QY 421 GCCCAGGACAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAACCAAGC 480

Db 457 GCCCAGGACAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAACCAAGC 516

QY 481 CCGTGGATCCCCCAATTTACAGTATGCTTTTCCAGCAAAAATCACCTTTATCTGTGGAG 540

Db	517	CCGTGGATCCCCAATTACAGTAGTCCTTTCAGGACAAAAATCACCTTTATCTGGTCATG	576
Qy	541	GAATATCAGCCTGGAGGGGACTTCTCTCACTTTTGAATAGATATGAGCCAGTTAGAT	600
Db	577	GAATATCAGCCTGGAGGGGACTTCTCTCACTTTTGAATAGATATGAGCCAGTTAGAT	636
Qy	601	GAATAACCTGATACAGTTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTCACTCG	660
Db	637	GAATAACCTGATACAGTTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTCACTCG	696
Qy	661	ATGGGATACGTGCATCGAGACATCAAGCCTTGAGAACATTTCTCGTTGACCCGACAGGACAC	720
Db	697	ATGGGATACGTGCATCGAGACATCAAGCCTTGAGAACATTTCTCGTTGACCCGACAGGACAC	756
Qy	721	ATCAAGCTGTTGGATTTTGGATCTGCGCGGAAATAATGAATTCAAACAAGATGGTGAATGCC	780
Db	757	ATCAAGCTGTTGGATTTTGGATCTGCGCGCGGAAATAATGAATTCAAACAAGATGGTGAATGCC	816
Qy	781	AAACTCCCGATTGGGACCCAGATTACATGCTTCCTGAAGTGTGACTGTGATGTAAGCGGG	840
Db	817	AAACTCCCGATTGGGACCCAGATTACATGCTTCCTGAAGTGTGACTGTGATGTAAGCGGG	876
Qy	841	GATGAAAAAGCCACTTACGSCCTTGGACTGTGACTGTGGTTCAGTGGGCGTGATTGCCAT	900
Db	877	GATGAAAAAGCCACTTACGSCCTTGGACTGTGACTGTGGTTCAGTGGGCGTGATTGCCAT	936
Qy	901	GAGATGATTTATGGGAGATCCCCCTTCGACAGGGGAACCTCTGCCAGAACCTTCAATAAC	960
Db	937	GAGATGATTTATGGGAGATCCCCCTTCGACAGGGGAACCTCTGCCAGAACCTTCAATAAC	996
Qy	961	ATTATGAATTTCCAGCCGGTTTTTGAATTTCCAGATGACCCCAAGTGACAGTGACTTT	1020
Db	997	ATTATGAATTTCCAGCCGGTTTTTGAATTTCCAGATGACCCCAAGTGACAGTGACTTT	1056
Qy	1021	CTTGATCTGATTTAAAGCTTGTGTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTT	1080
Db	1057	CTTGATCTGATTTAAAGCTTGTGTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTT	1116
Qy	1081	TGTGCGCATCCTTTCTCTCTAAATTTGACTGGAAACAATTCGTACTCTCTCTCCCCC	1140
Db	1117	TGTGCGCATCCTTTCTCTCTAAATTTGACTGGNAACAATTCGTACTCTCTCTCCCCC	1176
Qy	1141	TTGCTTCCCACCTTCAAGCTGACGATGACACCTTCCAAATTTTGATGAACACAGAGAAT	1200
Db	1177	TTGCTTCCCACCTTCAAGCTGACGATGACACCTTCCAAATTTTGATGAACACAGAGAAT	1236
Qy	1201	TCGTGGGTTTTATCCTCTCCGTGCGCACTGAGCCCCCTCAGGCTTCTCGGTGGAAGACTG	1260
Db	1237	TCGTGGGTTTTATCCTCTCCGTGCGCACTGAGCCCCCTCAGGCTTCTCGGTGGAAGACTG	1296
Qy	1261	CCGTTTTGTGGGTTTTCGTACAGAAAGCACTGGGATTTCTGGTAGACTGAGTCTGTT	1320
Db	1297	CCGTTTTGTGGGTTTTCGTACAGAAAGCACTGGGATTTCTGGTAGACTGAGTCTGTT	1356
Qy	1321	GTGTGCGGCTGGAATCCTCCCTGCCAGACTAGCTCCATGTGAAGAAAGAAATCTTCAATCAA	1380
Db	1357	GTGTGCGGCTGGAATCCTCCCTGCCAGACTAGCTCCATGTGAAGAAAGAAATCTTCAATCAA	1416
Qy	1381	AGCAAGAGCTACAAGACTCTCAGGACAAGTGTACAAAGATGCAGCAGGAAATGACCCCG	1440
Db	1417	AGCAAGAGCTACAAGACTCTCAGGACAAGTGTACAAAGATGCAGCAGGAAATGACCCCG	1476
Qy	1441	TTACATCGGAGAGTTCAGAGGTGGAGGCTGTGCTTAGTTCAGAAAGAGGTGGAGCTGAAG	1500
Db	1477	TTACATCGGAGAGTTCAGAGGTGGAGGCTGTGCTTAGTTCAGAAAGAGGTGGAGCTGAAG	1536
Qy	1501	GCCTCTCAGACTCAGAGATCCCTCCTCGAGCAGGACCTTCTACCTACATCACAGAAATGC	1560
Db	1537	GCCTCTCAGACTCAGAGATCCCTCCTCGAGCAGGACCTTCTACCTACATCACAGAAATGC	1596
Qy	1561	AGTAGCTTAAAGCGAAGTTTTGGAGCAAGCAGGATGGAGTGTCTCCAGGAGGATGACAA	1620
Db	1597	AGTAGCTTAAAGCGAAGTTTTGGAGCAAGCAGGATGGAGTGTCTCCAGGAGGATGACAA	1656

QY 2653 GAACTGGAGCAAGATTGCGGAGGTCAGTCTAGAGCACGAGAGCAGAAACTGAGGCTC 2712
Db 2737 GAACTGGAGCAAGATTGCGGAGGTCAGTCTAGAGCACGAGAGCAGAAACTGAGGCTC 2796
QY 2713 AAGCGCCAGCTCACAGAGCTACAGTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCC 2772
Db 2797 AAGCGCCAGCTCACAGAGCTACAGTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCC 2856
QY 2773 CTGCAAGCTGCAAGCGGCGCCCTGAGAGCCAGCTTCGCGAGGAGCGAGTCAAGTTGACAGCC 2832
Db 2857 CTGCAAGCTGCAAGCGGCGCCCTGAGAGCCAGCTTCGCGAGGAGCGAGTCAAGTTGACAGCC 2916
QY 2833 GAGACACAGCAGAGCTCAAGAGGAGATCCAGGAGCTCAAGGAGCTCAAGGAGCAGATGAAATC 2892
Db 2917 GAGACACAGCAGAGCTCAAGAGGAGATCCAGGAGCTCAAGGAGCAGATGAAATC 2976
QY 2893 CAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTAAGTAAATCACAGAGCTCGAGAGAGCAG 2952
Db 2977 CAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTAAGTAAATCACAGAGCTCGAGAGAGCAG 3036
QY 2953 CTAAACCAGCTGACCGAGAGCAACGCTGAACTCAACCAACCAAAATCTTCTACTTTGCCAAA 3012
Db 3037 CTAAACCAGCTGACCGAGAGCAACGCTGAACTCAACCAACCAAAATCTTCTACTTTGCCAAA 3096
QY 3013 CAACTCGATGAGGCTTCTGCGGCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGGAC 3072
Db 3097 CAACTCGATGAGGCTTCTGCGGCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGGAC 3156
QY 3073 CATCTCCGCGGAGATCACGAAACGAGAGATGCAAGCTTACAGCCAGAGCAAAACGATG 3132
Db 3157 CATCTCCGCGGAGATCACGAAACGAGAGATGCAAGCTTACAGCCAGAGCAAAACGATG 3216
QY 3133 GAGGCTCTCAAGCACAGTGCACCATGCTGAGGAGACAGGTCAATGGATTTGAGAGCCCTA 3192
Db 3217 GAGGCTCTCAAGCACAGTGCACCATGCTGAGGAGACAGGTCAATGGATTTGAGAGCCCTA 3276
QY 3193 AACGATGAGCTCTAGAAAAAGAGCGCAGTGGGAGCGCTGAGGAGCGTCTCGGGTGAT 3252
Db 3277 AACGATGAGCTCTAGAAAAAGAGCGCAGTGGGAGCGCTGAGGAGCGTCTCGGGTGAT 3336
QY 3253 GAGAAATCCAGTTTGAAGTGTGGGTTGAGAGCTGCGAGAGTGCAGAGAACTGCGACACCGAGAAA 3312
Db 3337 GAGAAATCCAGTTTGAAGTGTGGGTTGAGAGCTGCGAGAGTGCAGAGAACTGCGACACCGAGAAA 3396
QY 3313 CAGAGCAGGGCAGAGCCGATCAGCGGATCACCGAGTCTCGCAGGTGGTGAGCTGGCA 3372
Db 3397 CAGAGCAGGGCAGAGCCGATCAGCGGATCACCGGATCACCGAGTCTCGCAGGTGGTGAGCTGGCA 3456
QY 3373 GTGAAGGAGCAAGGCTGAGATTCTGCTCTGACAGAGGCTCTCAAAGAGCAGAAAGCTG 3432
Db 3457 GTGAAGGAGCAAGGCTGAGATTCTGCTCTGACAGAGGCTCTCAAAGAGCAGAAAGCTG 3516
QY 3433 AAGGCGAGAGCTCTCTGACAGCTCAATGACCTGGAGAGAGAGCAGTATGCTTGAA 3492
Db 3517 AAGGCGAGAGCTCTCTGACAGCTCAATGACCTGGAGAGAGAGCAGTATGCTTGAA 3576
QY 3493 ATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTT 3552
Db 3577 ATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTT 3636
QY 3553 CTGGAGAGCAGCCAAATTAAGCAGCAGATGAGACCTGCGACAAAATCAGATTTCCGT 3612
Db 3637 CTGGAGAGCAGCCAAATTAAGCAGCAGATGAGACCTGCGACAAAATCAGATTTCCGT 3696
QY 3613 CTGACTCAAGGAGCTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGT 3672
Db 3697 CTGACTCAAGGAGCTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGT 3756
QY 3673 GACTTGGAGTATCAGCTGGAAACATTCAGGTTCTCTATTTCTCATGAAAGGTGAAATG 3732
Db 3757 GACTTGGAGTATCAGCTGGAAACATTCAGGTTCTCTATTTCTCATGAAAGGTGAAATG 3816
QY 3733 GAAGGCACTATTCTCAACAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAA 3792

Db 3817 GAAGGCACTATTCTTCAACAAACCAAACTCATTGATTTCTGAAAGCCAAAATGAGCAA 3876
QY 3793 CCTCTAAAAGAAAAAGGGTTTATTAGTCGACGGAAGAGAGACCCTGCTTACCCACA 3852
Db 3877 CCTGCTAAAAGAAAA----- 3892
QY 3853 CAGGTTCTCTGCAATGACATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGT 3912
Db 3893 -AGGTTCTCTGCAATGACATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGT 3951
QY 3913 GCAGAGCTAGAGGAAGCCCTTCAGAGACCCGATCGAGTCCGCTCCGCCGCGAGGAA 3972
Db 3952 GCAGAGCTAGAGGAAGCCCTTCAGAGACCCGATCGAGTCCGCTCCGCCGCGAGGAA 4011
QY 3973 GCTGCCCAACGCAAAAGCAACGAGCCACCAACCCATCCAGCGCAGCAACGCGAGGAG 4032
Db 4012 GCTGCCCAACGCAAAAGCAACGAGCCACCAACCCATCCAGCGCAGCAACGCGAGGAG 4071
QY 4033 CAGATCCCATGTCGCGCATGTCGCGTCCGAGACACCCAGCCAGTGCATGAGGCTG 4092
Db 4072 CAGATCCCATGTCGCGCATGTCGCGTCCGAGACACCCAGCCAGTGCATGAGGCTG 4131
QY 4093 CTGCCCCCGCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCGCGCT 4152
Db 4132 CTGCCCCCGCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCGCGCT 4191
QY 4153 CTTAAGGAAACGATGACCAATATTTCTTCAACGATTCAGGTTCAACGTTAGGACTGAAATCGGA 4212
Db 4192 CTTAAGGAAACGATGACCAATATTTCTTCAACGATTCAGGTTCAACGTTAGGACTGAAATCGGA 4251
QY 4213 GCCAACAAGTGTGCTGTCTGGATACCGTGCACCTTGGAGCCAGGAGCATCCAAATGT 4272
Db 4252 GCCAACAAGTGTGCTGTCTGGATACCGTGCACCTTGGAGCCAGGAGCATCCAAATGT 4311
QY 4273 CTCGAATGTCAAGTGTGTCTACCCCAAGTGTCTCAAGTGTCTGCGAGCACTGCGGC 4332
Db 4312 CTCGAATGTCAAGTGTGTCTACCCCAAGTGTCTCAAGTGTCTGCGAGCACTGCGGC 4371
QY 4333 TTGCTCTGCAATATGACACACCTTCCAGCGGCTTCTGCGTGACAAATGAACTCC 4392
Db 4372 TTGCTCTGCAATATGACACACCTTCCAGCGGCTTCTGCGTGACAAATGAACTCC 4431
QY 4393 CCAGGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGACCTGGAAGGCTGGATGAAGTG 4452
Db 4432 CCAGGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGACCTGGAAGGCTGGATGAAGTG 4491
QY 4453 CCGAGGAATAACAAACGAGCAGCAGGCTGGAGAGGAGTACATTTGCTCGAGGGA 4512
Db 4492 CCGAGGAATAACAAACGAGCAGCAGGCTGGAGAGGAGTACATTTGCTCGAGGGA 4551
QY 4513 TCAAAAGTCTCATTTATGCAATGAAGCCAGAGAGCTGGACAGGCGCGGTGGAAGAA 4572
Db 4552 TCBAAGTCTCATTTATGCAATGAAGCCAGAGAGCTGGACAGAGGCGCGGTGGAAGAA 4611
QY 4573 TTTGAGTGTGCTTCCCGACGGGATGATCTATTATGCTGCGGCTTGTGCTTCCGAA 4632
Db 4612 TTTGAGTGTGCTTCCCGACGGGATGATCTATTATGCTGCGGCTTGTGCTTCCGAA 4671
QY 4633 CTCGCAAAATACAGCCAAAGCA----- 4653
Db 4672 CTCGCAAAATACAGCCAAAGCAGATGTCCCATACATCTGAAGATGGNATCTACCCGAC 4731
QY 4654 ----- 4653
Db 4732 ACCACCTGCTGGCCGGGAGAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAG 4791
QY 4654 ----- 4662
Db 4792 CGGTGGTCAACGCTTTAGAAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAAAAGCA 4851
QY 4663 GAAGCTGATGCTAAACTGCTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTA 4722

Db 4852 GAAGCTGATGCTAAACTGCTTGGAACTCCCTGCTGAAACTGGAAGGTGATGACCCGCTCTA 4911
QY 4723 GACATGAATGCGAGCGCCCTTCAGTGACACAGTGGTGTGGTGGCCACCGAGGAGG 4782
Db 4912 GACATGAATGCGAGCGCCCTTCAGTGACACAGTGGTGTGGTGGCCACCGAGGAGG 4971
QY 4783 CTCTACGCCCTGAATGCTTTGAAAACTCCCTAACCCATGTCCAGGAAATTTGAGCAGTC 4842
Db 4972 CTCTACGCCCTGAATGCTTTGAAAACTCCCTAACCCATGTCCAGGAAATTTGAGCAGTC 5031
QY 4843 TTCCAAATTTATATATCAAGGACCTGGAGAACTACTCATGATAGCAGGAGAGCGG 4902
Db 5032 TTCCAAATTTATATATCAAGGACCTGGAGAACTACTCATGATAGCAGGAGAGCGG 5091
QY 4903 GCACGTGTCTTTGGAGCGTGAAGAAAGTGAACAGCTCCCTGGCCAGTCCCACTGCGCT 4962
Db 5092 GCACGTGTCTTTGGAGCGTGAAGAAAGTGAACAGCTCCCTGGCCAGTCCCACTGCGCT 5151
QY 4963 GCCAGCCCGACATCTCACCCACATTTTGAAGCTGTCAAGGCTGGCCACTTGTGGG 5022
Db 5152 GCCAGCCCGACATCTCACCCACATTTTGAAGCTGTCAAGGCTGGCCACTTGTGGG 5211
QY 5023 GCAGGCAAGATTTGAGAACGGCTCTGCATCTGTGACGACCATGCCAGCAAGTCTGTCAT 5082
Db 5212 GCAGGCAAGATTTGAGAACGGCTCTGCATCTGTGACGACCATGCCAGCAAGTCTGTCAT 5271
QY 5083 CTCGCTACCAACGAAACCTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAG 5142
Db 5272 CTCGCTACCAACGAAACCTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAG 5331
QY 5143 CCTCGAGCTGTATCCACTTCACAAATACAGTATCTCATTTGGAACCAATTAATCTAC 5202
Db 5332 CCTCGAGCTGTATCCACTTCACAAATACAGTATCTCATTTGGAACCAATTAATCTAC 5391
QY 5203 GAAATCGACATGAAGCAGTACAGCTCGAGGAAATCTCTGATAAGAAATGACCAATCTCTTG 5262
Db 5392 GAAATCGACATGAAGCAGTACAGCTCGAGGAAATCTCTGATAAGAAATGACCAATCTCTTG 5451
QY 5263 GCACCTGTCTGTTGGCGCTCTTCCAAACAGCTTCCCTGTCTCAATCTGCAGGTGAC 5322
Db 5452 GCACCTGTCTGTTGGCGCTCTTCCAAACAGCTTCCCTGTCTCAATCTGCAGGTGAC 5511
QY 5323 AGCGAGGCGAGGAGGAGTACTTGTCTGTTTCCAGAAATTTGAGTGTCTGTGGAT 5382
Db 5512 AGCGAGGCGAGGAGGAGTACTTGTCTGTTTCCAGAAATTTGAGTGTCTGTGGAT 5571
QY 5383 TCTTACGGAAGAGTGTAGCGCAGACAGATCTCAAGTGAGTGCCTTACCTTTGGCTTTT 5442
Db 5572 TCTTACGGAAGAGTGTAGCGCAGACAGATCTCAAGTGAGTGCCTTACCTTTGGCTTTT 5631
QY 5443 GCCTACAGAGAAACCTTATCTGTTTGTGACCACTTCAACTCTCAAGTGAATTTGAGATC 5502
Db 5632 GCCTACAGAGAAACCTTATCTGTTTGTGACCACTTCAACTCTCAAGTGAATTTGAGATC 5691
QY 5503 CAGGCAAGCTCTCAGCAGGACCCCTGCGGAGCGTACTCGACATCCCGAACCCGCGC 5562
Db 5692 CAGGCAAGCTCTCAGCAGGACCCCTGCGGAGCGTACTCGACATCCCGAACCCGCGC 5751
QY 5563 TAGCTGGCCCTGCGCATTTCTCAGAGGAGTATTACTTGGCGTCTCATACAGGATAA 5622
Db 5752 TAGCTGGCCCTGCGCATTTCTCAGAGGAGTATTACTTGGCGTCTCATACAGGATAA 5811
QY 5623 TTAAGGCTCATTTGCTGCAAGGAAACCTCGTGAAGGAGTCCGGCAGCTGAACACCCAGG 5682
Db 5812 TTAAGGCTCATTTGCTGCAAGGAAACCTCGTGAAGGAGTCCGGCAGCTGAACACCCAGG 5871
QY 5683 GGCCGCTCAGCTCCGCGACAGCCCAACAGGAGGCGCCACCCAGCTACACAGGAC 5742
Db 5872 GGCCGCTCAGCTCCGCGACAGCCCAACAGGAGGCGCCACCCAGCTACACAGGAC 5931
QY 5743 ATCACAAGCGCGTGGCTCTCAGCCAGGCGCCGCCGAGGCGCCAGCCACCCGCGAGAG 5802
Db 5932 ATCACAAGCGCGTGGCTCTCAGCCAGGCGCCGCCGAGGCGCCAGCCACCCGCGAGAG 5991

QY 5803 CCAAGCACACCCACCGCTACCGAGGCGGCGGACCGAGCTGCGCAGGACCAAGTCTCCT 5862
Db 5992 CCAAGCACACCCACCGCTACCGAGGCGGCGGACCGAGCTGCGCAGGACCAAGTCTCCT 6051
QY 5863 GGCGCCGCCCTTGGAGCGAGAGAAAGTCCCGCGGATGCTCAGCAGCGGAGAGAGCGG 5922
Db 6052 GGCGCCGCCCTTGGAGCGAGAGAAAGTCCCGCGGATGCTCAGCAGCGGAGAGAGCGG 6111
QY 5923 TCCCGCGGAGCGCTTTTGAAGACAGCAGAGCGGCGGCTGCTGCGGAGCGCTGAGG 5982
Db 6112 TCCCGCGGAGCGCTTTTGAAGACAGCAGAGCGGCGGCTGCTGCGGAGCGCTGAGG 6171
QY 5983 ACCCGCTGTCCCGAGGTGAACAGGAGGCGGAGAGTGCCTCTCAAGTTTTCACGGTT 6042
Db 6172 ACCCGCTGTCCCGAGGTGAACAGGAGTGCCTGCGGACCGAGTTCAGTATAATCTCAGCCAGA 6231
QY 6043 AACACTGTCACTATTAT 6060
Db 6232 AAAACCAACTCCTCATCT 6249

RESULT 3

AAL55215

ID AAL55215 standard; DNA; 8603 BP.

AC AAL55215;

XX

DT 01-MAY-2003 (first entry)

XX

DE Human CR1K related DNA sequence, SEQ ID No 4.

XX

KW Anorectic; hypotensive; cardiac; antilipemic; cerebroprotective;
KW angiot; osteopathic; antiarthritic; cytostatic; antidepressant;
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;
KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;
KW central nervous system disorder; chronic obstructive pulmonary disease;
KW diabetes; pain; ds.

XX

OS Homo sapiens.

XX

PN WO2003004523-A1.

XX

PD 16-JAN-2003.

XX

PF 28-JUN-2002; 2002WO-EP007156.

XX

PR 02-JUL-2001; 2001US-0301841P.

XX

PR 11-DEC-2001; 2001US-0338651P.

XX

PR 25-APR-2002; 2002US-0375014P.

XX

PA (FARB) BAYER AG.

XX

PI Zhu Z;

XX

XX WPI; 2003-221576/21.

XX

PT New human citron rho/rac-interacting kinase (CR1K) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.

XX

PS Disclosure; Fig 4; 237pp; English.

XX

CC The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CR1K)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CR1K dysfunction, such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery

disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of cancer including endometrial, breast, prostate and colon cancer), anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood disorders, anxiety disorders, Parkinson's disease or Alzheimer's disease), chronic obstructive pulmonary disease, or diabetes. These can also be used to treat pain associated with the disorders. The human CR1K polypeptide is also useful in diagnostic assays or in genetic testing. The expression vector or the reagent is useful in preparing a medicament for modulating the activity of a human CR1K in a disease, e.g. obesity, a central nervous system disorder, or chronic obstructive pulmonary disease. The fusion protein is useful for generating antibodies against a CR1K polypeptide and for use in various assay systems. The methods are useful in producing and detecting the polynucleotide and polypeptide and in screening for agents that modulate the activity of the human CR1K polypeptide. This polynucleotide sequence represents a DNA sequence relating to the human CR1K protein of the invention.

```

X
Q Sequence 8603 BP; 2305 A; 2206 C; 2215 G; 1877 T; 0 U; 0 Other;
Query Match 86.2%; Score 5668; DB 8; Length 8603;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5962; Conservative 0; Mismatches 35; Indels 243; Gaps 3

```

19	ATGTTGAAGTTC	AAATATATGAGCGCGGAATCC	TTTGGATGCTGTGGTCTGCTGCTGAACCCCAATT	78
1	ATGTTGAAGTTC	AAATATATGAGCGCGGAATCC	TTTGGATGCTGTGGTCTGCTGCTGAACCCCAATT	60
79	GCAGCGCGGCTC	TCAGGCTGAATCTGTTC	TTCCAGGGGAAACCAACCTTTATGACTCAA	138
61	GCAGCGCGGCTC	CNAGGCTGAATCTGTTC	TTCCAGGGGNAACCAACCTTTATGACTCAA	120
139	CAGCAGATGCTC	CTCTTTCCCGAGAAGGGAATT	AGATGCCTCTTTGTTCTCTTTGAA	198
121	CAGCAGATGCTC	CTCTTTCCCGAGAAGGGAATT	AGATGCCTCTTTGTTCTCTTTGAA	180
199	GAA TCAGTCAGCTG	CTCTGATGAGATTAA	GACGACGTTGTCGCGGAAGTAT	258
181	GAA TCAGTCAGCTG	CTCTGATGAGATTAA	GACGACGTTGTCGCGGAAGTAT	240
259	TCGGACACATAG	CTGAGTTACAGGAGCTCC	AGCCTTCGGCAAAAGGACTTCGAACTCAGA	318
241	TCGGACACCATAG	CTGAGTTACAGGAGCTCC	AGCCTTCGGCAAAAGGACTTCGAACTCAGA	300
319	AGTCTTGTAAGT	TGTGGTCACTTTGCTGAAGTGC	AGTGGTGAAGAGAAACCAACCGGG	378
301	AGTCTTGTAAGT	TGTGGTCACTTTGCTGAAGTGC	AGTGGTGAAGAGAAACCAACCGGG	360
379	GACATCTATGCT	ATGAACTGATGAAGAAGAAGGCTTT	ATTCGCCACGAGAGCAGGTTTCA	438
361	GACATCTATGCT	ATGAACTGATGAAGAAGAAGGCTTT	ATTCGCCACGAGAGCAGGTTTCA	420
439	TTTTTTGAGGAAG	AGCGGAACATATTATCTCGAAGCACA	AGCCGCTGGATCCCCCAATTA	498
421	TTTTTTGAGGAAG	AGCGGAACATATTATCTCGAAGCACA	AGCCGCTGGATCCCCCAATTA	480
499	CAGTATGCCTTT	CAGACACAAAATCACCTTTATCTGAT	CGAGGAATATCAGCCTGGAGGG	558
481	CAGTATGCCTTT	CAGACACAAAATCACCTTTATCTGAT	CGAGGAATATCAGCCTGGAGGG	540
559	GACTTCGTCTC	ATTTGAAATGATATGAGGACCA	GTTAGATTAGATAACCTGATCAGTTT	618
541	GACTTCGTCTC	ATTTGAAATGATATGAGGACCA	GTTAGATTAGATAACCTGATCAGTTT	600
619	TACCTAGCTGAG	CTGATTTGGCTGTTCTACAGGTTTCAT	CTGATGGGATACGTGATCGA	678
601	TACCTAGCTGAG	CTGATTTGGCTGTTCTACAGGTTTCAT	CTGATGGGATACGTGATCGA	660
679	GACATCAAGCCT	CAGAACATTTCTGTTTGACCGCAC	AGGACACATCAAGCTGGTGATTTT	738
661	GACATCAAGCCT	CAGAACATTTCTGTTTGACCGCAC	AGGACACATCAAGCTGGTGATTTT	720
739	GGATCTCGCGCG	AAAAATGAATTC	CAAAACAAGATGGTGAATGCCAAACCTCCGATTTGGGACC	798

QY 1879 GTGGGAGATATGCGAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938
DB 1861 GTGGGAGATATGCGAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
QY 1939 CTCGACAGAGAACTGGAGAGAGCTG----- 1963
DB 1921 CTCGAGAGAACTGGAGAGAGCTGTTAAAGCCAGCAGCGAGGCCACCGAGCTGCTGCGAG 1980
QY 1964 -----CAAGAGCGAGCGAGCGAGGAGCTGGAGAGCTGGAGAGCTGCAGAACCGAGAG 2010
DB 1981 AATATCCGCGCAGGCAAGAGGCGAGCGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2040
QY 2011 GATTCCTTCGAGGCGATCAGAAAGAGCTGGTGGAGAGCTGAGGAGCGCGCCATCTCTG 2070
DB 2041 GATTCCTTCGAGGCGATCAGAAAGAGCTGGTGGAGAGCTGAGGAGCGCGCCATCTCTG 2100
QY 2071 GAGAACCAAGGTAAAGAGACTAGAGACCATGGAGCGCTAGAGAAACAGACTGGAAGGATGAC 2130
DB 2101 GAGAACCAAGGTAAAGAGACTAGAGACCATGGAGCGCTAGAGAAACAGACTGGAAGGATGAC 2160
QY 2131 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAGAG 2190
DB 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAGAG 2220
QY 2191 AAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAAACAGAGAGAG 2250
DB 2221 AAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAAACAGAGAGAG 2280
QY 2251 CAGCACTATGAGGAAAGATTAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2310
DB 2281 CAGCACTATGAGGAAAGATTAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340
QY 2311 AAGGAGACCTGGAGAACATGATGAGAGACAGAGAGGAGGCCCATGAGAGGSCAAA 2370
DB 2341 AAGGAGACCTGGAGAACATGATGAGAGACAGAGAGGAGGCCCATGAGAGGSCAAA 2400
QY 2371 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTCGAAAGTTCAGATCCCTGGAA 2430
DB 2401 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTCGAAAGTTCAGATCCCTGGAA 2460
QY 2431 CAGAGGATGTGGAACCTGCTGAGGCCAATAACTTGCACCAATAGCAGCTTTTACC 2490
DB 2461 CAGAGGATGTGGAACCTGCTGAGGCCAATAACTTGCACCAATAGCAGCTTTTACC 2520
QY 2491 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550
DB 2521 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
QY 2551 CTGGAGACAGAGCTGGAGAGTTGGAGGCCGAGAACCGGAAACTGGAGGAGCAGCTGGAG 2610
DB 2581 CTGGAGACAGAGCTGGAGAGTTGGAGGCCGAGAACCGGAAACTGGAGGAGCAGCTGGAG 2640
QY 2611 AAGATCAGGCCACCAAGACAGTGCAGAAATCGGCTGCTGAACTGGAGACAGAGATTG 2670
DB 2641 AAGATCAGGCCACCAAGACAGTGCAGAAATCGGCTGCTGAACTGGAGACAGAGATTG 2700
QY 2671 CGGAGGTCAAGTCTAGAGACAGAGGAGCAGAACTGGAGCTCAAGGCCAGCTCAGAG 2730
DB 2701 CGGAGGTCAAGTCTAGAGACAGAGGAGCAGAACTGGAGCTCAAGGCCAGCTCAGAG 2760
QY 2731 CTACAGCTCTCCCTCAGAGGCGGAGTGCAGTTCACAGCCCTGCAGGCTCAGCGGCG 2790
DB 2761 CTACAGCTCTCCCTCAGAGGCGGAGTGCAGTTCACAGCCCTGCAGGCTCAGCGGCG 2820
QY 2791 GCCTTGAGAGCAGCTTGGCAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGCT 2850
DB 2821 GCCTTGAGAGCAGCTTGGCAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGCT 2880
QY 2851 GAAGAGGATCCAGGCACTCAGGCGACATAGAGATGAAATCCAGGCGAAATTTGATGCT 2910
DB 2881 GAAGAGGATCCAGGCACTCAGGCGACATAGAGATGAAATCCAGGCGAAATTTGATGCT 2940

QY 2911 CTTGTAACAGCTGTACTGTAATCAGACCTGGAGGAGAGCTAAACCAGCTGACCGAG 2970
DB 2941 CTTGTAACAGCTGTACTGTAATCAGACCTGGAGGAGAGCTAAACCAGCTGACCGAG 3000
QY 2971 GACAAACGCTGAACCTCAACCAACAAAACCTTCTATTGTTGCCAAACAACTCGATGAGGCTTCT 3030
DB 3001 GACAAACGCTGAACCTCAACCAACAAAACCTTCTATTGTTGCCAAACAACTCGATGAGGCTTCT 3060
QY 3031 GCGGCCAAGCAGAGATTGTACAACTCGGAAGTGAAGTGAACCATCTCCGCGCGGAGATC 3090
DB 3061 GCGGCCAAGCAGAGATTGTACAACTCGGAAGTGAAGTGAACCATCTCCGCGCGGAGATC 3120
QY 3091 ACGGAAACGAGAGATGACAGCTTACCAGCAGAGCAACAGATGAGGCTCTGAAAGCCACG 3150
DB 3121 ACGGAAACGAGAGATGACAGCTTACCAGCAGAGCAACAGATGAGGCTCTGAAAGCCACG 3180
QY 3151 TGCACCAATGCTGGAGGAACAGGTCATGATTTGGAGGCCCTAAACGATGAGTGTCTAGAA 3210
DB 3181 TGCACCAATGCTGGAGGAACAGGTCATGATTTGGAGGCCCTAAACGATGAGTGTCTAGAA 3240
QY 3211 AAGAGCGGCGAGTGGAGGCGCTGGAGGAGCGTCTGGGTGATGAGAAATCCCAAGTTTGG 3270
DB 3241 AAGAGCGGCGAGTGGAGGCGCTGGAGGAGCGTCTGGGTGATGAGAAATCCCAAGTTTGG 3300
QY 3271 TGTCCGGTTCCAGAGCTGCGAGAGAAATCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3330
DB 3301 TGTCCGGTTCCAGAGCTGCGAGAGAAATCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3360
QY 3331 GATCAGCGGATCACCGAGTCTGCCAGGCTGGTGGAGCTGCAGTGGAGGAGCAGAGGCT 3390
DB 3361 GATCAGCGGATCACCGAGTCTGCCAGGCTGGTGGAGCTGCAGTGGAGGAGCAGAGGCT 3420
QY 3391 GAGATTCTCGCTCTGCGAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3450
DB 3421 GAGATTCTCGCTCTGCGAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3480
QY 3451 GACAACTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCAGAGCTTA 3510
DB 3481 GACAACTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCAGAGCTTA 3540
QY 3511 CAGCAGAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTCTGGAAGAGCAGAGCAAA 3570
DB 3541 CAGCAGAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTCTGGAAGAGCAGAGCAAA 3600
QY 3571 TTAACAGCAGAGATGAGACCTGCGAGAAAAATCAATTTTCGCTGTGACTCAAGGACTGCAA 3630
DB 3601 TTAACAGCAGAGATGAGACCTGCGAGAAAAATCAATTTTCGCTGTGACTCAAGGACTGCAA 3660
QY 3631 GAAGCTCTAGATCGGCTGATCTGAGCAGAGAGAGTGAATGAGTGGAGTATCAGCTG 3690
DB 3661 GAAGCTCTAGATCGGCTGATCTGAGCAGAGAGAGTGAATGAGTGGAGTATCAGCTG 3720
QY 3691 GAAACCAATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGCAGCTATTTCTCAA 3750
DB 3721 GAAACCAATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGCAGCTATTTCTCAA 3780
QY 3751 CAAACCAATTCAGTTTCTGCAAGCCAAATGGAACCACTGCTGAAAGAGAAAG 3810
DB 3781 CAAACCAATTCAGTTTCTGCAAGCCAAATGGAACCACTGCTGAAAGAGAAAG -- 3838
QY 3811 GGTATTATTAGTCGACGGAAGAGGACCTGCTTTACCCACAGAGTTCTCTGCAAGTAC 3870
DB 3839 -----AGGTTCTCTGCAAGTAC 3855
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAAAGCTGCTGTGCAAGCTTAGAGGAGCC 3930
DB 3856 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGAAAGCTGCTGTGCAAGCTTAGAGGAGCC 3915
QY 3931 CTTGAGAGAGCCGCACTGAGCTCCGCTCCGCGGAGGAGAACTGCCACCCGCAAGCA 3990
DB 3916 CTTGAGAGAGCCGCACTGAGCTCCGCTCCGCGGAGGAGAACTGCCACCCGCAAGCA 3975
QY 3991 ACGGACCAACCAACCATCCACGCGCAGCCACCGGAGGAGCAGATCGCCATGTGCGCC 4050

Db 3976 ACAGACCCACCCATCCAGCCAGCCACCGGAGGACGAGATGCCATGTCCGCC 4035
Qy 4051 ATCTGTCGGTGCAGAGACACAGCCCATGAGCTCTGTCGCCGCCCATCCAGC 4110
Db 4036 ATCTGTCGGTGCAGAGACACAGCCCATGAGCTCTGTCGCCGCCCATCCAGC 4095
Qy 4111 CCAGAAAGGAGTCTCAACTCCAGAGAAATTAAGTCGGCTCTTAAGGAACGATGCAC 4170
Db 4096 CCAGAAAGGAGTCTCAACTCCAGAGAAATTAAGTCGGCTCTTAAGGAACGATGCAC 4155
Qy 4171 CACAAATATCTCTCACCGATTCAACCTAGGACTGAAATCGCAGCCACAAAGTGTCTGTG 4230
Db 4156 CACAAATATCTCTCACCGATTCAACCTAGGACTGAAATCGCAGCCACAAAGTGTCTGTG 4215
Qy 4231 TGTCTGGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCTCGAAATGTCAGGTGATG 4290
Db 4216 TGTCTGGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCTCGAAATGTCAGGTGATG 4275
Qy 4291 TGTCTACCCCAAGTGTCTCACAGTGTTCGACGCACCTCGGCTTGCCTGCTGAAATATGCC 4350
Db 4276 TGTCTACCCCAAGTGTCTCACAGTGTTCGACGCACCTCGGCTTGCCTGCTGAAATATGCC 4335
Qy 4351 ACACATCTCACCGAGCCTTTGCGTGACAAATGAACTCCCGAGGTCTCCAGACCAAG 4410
Db 4336 ACACATCTCACCGAGCCTTTGCGGTGACAAATGAACTCCCGAGGTCTCCAGACCAAG 4395
Qy 4411 GAGCCACAGCAGCTTCACCTGGAAGGTGGATGAGGTGCCAGGAATAACAAAGCA 4470
Db 4396 GAGCCACAGCAGCTTCACCTGGAAGGTGGATGAGGTGCCAGGAATAACAAAGCA 4455
Qy 4471 GGACAGCAAGGTGGGACAGGAAGTACATTTGCTGGAGGGATCAAAAGTCTCTATTAT 4530
Db 4456 GGACAGCAAGGTGGGACAGGAAGTACATTTGCTGGAGGGATCAAAAGTCTCTATTAT 4515
Qy 4531 GACATGAAAGCAGAGAGCTGGAAGGCGGTGGAGAAATTTGAGCTGTGCTTCCC 4590
Db 4516 GACATGAAAGCAGAGAGCTGGAAGGCGGTGGAGAAATTTGAGCTGTGCTTCCC 4575
Qy 4591 GACGGGATGTATCTATTATCATGTCGCTGCTCCGAACTCGCAATACAGCCAAA 4650
Db 4576 GACGGGATGTATCTATTATCATGTCGCTGCTCCGAACTCGCAATACAGCCAAA 4635
Qy 4651 GCA----- 4653
Db 4636 GCAGATGTCCCATACATACTGAAGTGAATCTCACCGGCACACCACTGTGGCCCGG 4695
Qy 4654 ----- 4653
Db 4696 AGAACCTCTACTTGTAGCTCCCAGCTTCCCTGACAAACAGCGCTGGGTACCCGCTTA 4755
Qy 4654 -----GAAAAAGCAGAAAGCTGATGCTAAACTG 4680
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAGAAAGCAGAGCTGATGCTAAACTG 4815
Qy 4681 CTTGAAACTCCCTGCTGAACTGAAGGTGATGACCGCTAGACATGAATGACGCTG 4740
Db 4816 CTTGAAACTCCCTGCTGAACTGAAGGTGATGACCGCTAGACATGAATGACGCTG 4875
Qy 4741 CCGTTCACTGACAGGTGCTGTTGGGCACCGCAGGAAGGCTCTACGCGCTGAATGTC 4800
Db 4876 CCGTTCACTGACAGGTGCTGTTGGGCACCGCAGGAAGGCTCTACGCGCTGAATGTC 4935
Qy 4801 TTGAAAACTCCCTAACCCATCTCCAGGAATTGAGCAGTCTTCCAAATTTATATTATC 4860
Db 4936 TTGAAAACTCCCTAACCCATCTCCAGGAATTGAGCAGTCTTCCAAATTTATATTATC 4995
Qy 4861 AAGGACCTGGAGAGTCTCATGATAGCAGAGAGAGCGGCACTGTCTTGTGGAC 4920
Db 4996 AAGGACCTGGAGAGTCTCATGATAGCAGAGAGAGCGGCACTGTCTTGTGGAC 5055
Qy 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4980

Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCAGCTGCCTGCCAGCCGACATCTCA 5115
Qy 4981 CCACAAATTTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGCAAGATTGAGAAC 5040
Db 5116 CCACAAATTTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGCAAGATTGAGAAC 5175
Qy 5041 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTCTATTCTCGCTACAAACGAAAC 5100
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTCTATTCTCGCTACAAACGAAAC 5235
Qy 5101 CTCAGCAAAATCTGCATTCGGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160
Db 5236 CTCAGCAAAATCTGCATTCGGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
Qy 5161 TTCACCAATTTACAGTATCTCTATTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
Db 5296 TTCACCAATTTACAGTATCTCTATTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
Qy 5221 TACACGCTCAGGAATTCCTGGATAAGAATGACCATTCCTTTGGCACCTGCTGTGTTGCC 5280
Db 5356 TACACGCTCAGGAATTCCTGGATAAGAATGACCATTCCTTTGGCACCTGCTGTGTTGCC 5415
Qy 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAAACAGCGGAGGCGAGCAGAG 5340
Db 5416 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAAACAGCGGAGGCGAGCAGAG 5475
Qy 5341 GAGTACTTGTCTGTTTCCACGAATTTGGAGTGTCTCTGTGATTTCTTACGGAAGACGTAGC 5400
Db 5476 GAGTACTTGTCTGTTTCCACGAATTTGGAGTGTCTCTGTGATTTCTTACGGAAGACGTAGC 5535
Qy 5401 CGCACAGACATCTCAAGTGGATGCTTACCTTTGGCCTTTGCCCTACAGAGAACCTTAT 5460
Db 5536 CGCACAGACATCTCAAGTGGATGCTTACCTTTGGCCTTTGCCCTACAGAGAACCTTAT 5595
Qy 5461 CTGTTTGTGACCCACATTCACCTCAGTGAATTTAGATTCAGGACAGCCTCTCAGCA 5520
Db 5596 CTGTTTGTGACCCACATTCACCTCAGTGAATTTAGATTCAGGACAGCCTCTCAGCA 5655
Qy 5521 GGGACCCCTCCCGAGCGTACCTGGAATCCCGAAACCCGCGTACCTGGGCCCTGCAATT 5580
Db 5656 GGGACCCCTCCCGAGCGTACCTGGAATCCCGAAACCCGCGTACCTGGGCCCTGCAATT 5715
Qy 5581 TCCTCAGGACGATTTACTTTGGCTCTCATACAGGATTAATTAAGGTTCATTTGCTGC 5640
Db 5716 TCCTCAGGACGATTTACTTTGGCTCTCATACAGGATTAATTAAGGTTCATTTGCTGC 5775
Qy 5641 AAGGGAAACCTCGTGAAGAGTCCGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5700
Db 5776 AAGGGAAACCTCGTGAAGAGTCCGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5835
Qy 5701 AGCAGCCCCAACAGGAGGCCACCCACGTA CAACGAGCAATCAACAGCGGTGGCC 5760
Db 5836 AGCAGCCCCAACAGGAGGCCACCCACGTA CAACGAGCAATCAACAGCGGTGGCC 5895
Qy 5761 TCAGCCACCGCGCCCGAGGCCACCCAGCCAGCGAGGCAAGACACACCCACCGC 5820
Db 5896 TCAGCCACCGCGCCCGAGGCCACCCAGCCAGCGAGGCAAGACACACCCACCGC 5955
Qy 5821 TACCGCAGGGGCGGACCGAGCTCGCAGGGAAGTCTCTCGCGCCGCCCTTGGAGCGA 5880
Db 5956 TACCGCAGGGGCGGACCGAGCTCGCAGGGAAGTCTCTCGCGCCGCCCTTGGAGCGA 6015
Qy 5881 GAGAAAGTCCCGCCCGGATGCTCAGCACGCGAGAGAGCGGTCCCGGAGGCTGTTT 5940
Db 6016 GAGAAAGTCCCGCCCGGATGCTCAGCACGCGAGAGAGCGGTCCCGGAGGCTGTTT 6075
Qy 5941 GAAGACAGCAGCAGGGGCCCGCTGCTGCGGAGCGGTGAGGACCCCGCTGTGCCAGGTG 6000
Db 6076 GAAGACAGCAGCAGGGGCCCGCTGCTGCGGAGCGGTGAGGACCCCGCTGTGCCAGGTG 6135
Qy 6001 AACAGGGAAGAGGAGAGTGCCTCTCAAGTTTTTACCGTTAAACTGTCACTTAT 6060
Db 6136 AACAGGTCTGGACCAAGTCTTCAGTATAAATCTCAGCCAGAAAAAACCACTCTCTCATCT 6195

Db	1021	TTGTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTTGCTGCCACATCTTCTTC	1080
Qy	1099	TCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCCCTCCCTTCGTTCCACCCCTCAAG	1158
Db	1081	TCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCCCTCCCTTCGTTCCACCCCTCAAG	1140
Qy	1159	TCTGACGATGACACCTCCAAATTTTGATGACCAAGAGAAATTCGTGGGTTTTCATCCTCT	1218
Db	1141	TCTGACGATGACACCTCCAAATTTTGATGACCAAGAGAAATTCGTGGGTTTTCATCCTCT	1200
Qy	1219	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCGGTTTGTGGGGTTTTCG	1278
Db	1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCGGTTTGTGGGGTTTTCG	1260
Qy	1279	TACAGCAAGCACTGGGATCTTGTAGATCTGAGTCTGTGTGTGCGGTCTGGAATCC	1338
Db	1261	TACAGCAAGCACTGGGATCTTGTAGATCTGAGTCTGTGTGTGCGGTCTGGAATCC	1320
Qy	1339	CTGTCCAAAGTACTAGCTCCATCGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1398
Db	1321	CTGTCCAAAGTACTAGCTCCATCGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1380
Qy	1399	TCTCAGCACAAAGTGTCACAGATGGAGCAGGAATAATGACCCGGTTACATCGGAGGTGTC	1458
Db	1381	TCTCAGCACAAAGTGTCACAGATGGAGCAGGAATAATGACCCGGTTACATCGGAGGTGTC	1440
Qy	1459	GAGGTGAGGCTGTGCTTACTCAGAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA	1518
Db	1441	GAGGTGAGGCTGTGCTTACTCAGAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA	1500
Qy	1519	TCCTCTCTGAGCAGGACCTTGTCTACTACATCACAGAATGCAGTAGCTTAAAGCGAAGT	1578
Db	1501	TCCTCTCTGAGCAGGACCTTGTCTACTACATCACAGAATGCAGTAGCTTAAAGCGAAGT	1560
Qy	1579	TTGAGCAAGCACGGATGGAGGTGTCCAGGAGGATGACAAAGCACTGAGGCTTCTCCAT	1638
Db	1561	TTGAGCAAGCACGGATGGAGGTGTCCAGGAGGATGACAAAGCACTGAGGCTTCTCCAT	1620
Qy	1639	GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGCAGGAGTACCAGGCTCAA	1698
Db	1621	GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGCAGGAGTACCAGGCTCAA	1680
Qy	1699	GTGGAAGAAATAGGTTTGATGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAAG	1758
Db	1681	GTGGAAGAAATAGGTTTGATGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAAG	1740
Qy	1759	CGGAGTCAATCTCTACGAATCTGAGCTCAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG	1818
Db	1741	CGGAGTCAATCTCTACGAATCTGAGCTCAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG	1800
Qy	1819	CGGAAAGCGACAGAATCTCAGCATAAATCTGTGAAGGCTTAAGGATCAAGGGAAGCCTGAA	1878
Db	1801	CGGAAAGCGACAGAATCTCAGCATAAATCTGTGAAGGCTTAAGGATCAAGGGAAGCCTGAA	1860
Qy	1879	GTGGGAGAAATGCGAACTGGAGAGATCAATGCTCAGCAGCAGCTCAAAATTCAGGAG	1938
Db	1861	GTGGGAGAAATGCGAACTGGAGAGATCAATGCTCAGCAGCAGCTCAAAATTCAGGAG	1920
Qy	1939	CTCCAAAGAGAACTGGAGAAAGGCT-----	1962
Db	1921	CTCCAAAGAGAACTGGAGAAAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAG	1980
Qy	1963	-----GCAAGAGCCGAGCCGAGAGGAGCTGGAGAAAGCTTGAGAAACCGAGAG	2010
Db	1981	AATATCCGCCAGGCAAGAGGCGAGCCGAGAGGAGCTGGAGAGCTGCGAAGCCGAGAG	2040
Qy	2011	GATTCTTTGTAAGGCATCAGAAAAGAGCTGGTGAAGCTGAGGAAACCGCCGCAATCTCTCG	2070
Db	2041	GATTCTTTGTAAGGCATCAGAAAAGAGCTGGTGAAGCTGAGGAAACCGCCGCAATCTCTCG	2100
Qy	2071	GAGAACAGGTPAAGAGACTTAGACCATCGAGCGTAGAGAAAACAGACTGAAGGATGAC	2130
Db	2101	GAGAACAGGTPAAGAGACTTAGACCATCGAGCGTAGAGAAAACAGACTGAAGGATGAC	2160

QY	2131	ATCAGACAAAATCCCAACAGATCCAGCAGATGCTGATAAAATTTCTGGAGACTCGAAGAG	21190
DB	2161	ATCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG	2220
QY	2191	AAAATCGGAGGCCCAAGTCTCAGCCACGACCTAGAAGTGCACCTGAAACAGAAAGAG	2250
DB	2221	AAAATCGGAGGCCCAAGTCTCAGCCACGACCTAGAAGTGCACCTGAAACAGAAAGAG	2280
QY	2251	CAGCACTATGAGGAAAAGATTAAAGTCTTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2310
DB	2281	CAGCACTATGAGGAAAAGATTAAAGTCTTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2340
QY	2311	AAGGAGACACTGGAGAACATGATGCAGAGACAGAGGAGGAGGCCCATGAGAAAGGCCAAA	2370
DB	2341	AAGGAGACACTGGAGAACATGATGCAGAGACAGAGGAGGAGGCCCATGAGAAAGGCCAAA	2400
QY	2371	ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCGAGATCCAGATCAGATCCCTGGAA	2430
DB	2401	ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCGAGATCCAGATCAGATCCCTGGAA	2460
QY	2431	CAGAGGATTGTGGAACCTGTCTGAAGCCAAATAAATTCAGACAAATAGCAGTCTTTTATCC	2490
DB	2461	CAGAGGATTGTGGAACCTGTCTGAAGCCAAATAAATTCAGACAAATAGCAGTCTTTTATCC	2520
QY	2491	CAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC	2550
DB	2521	CAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC	2580
QY	2551	CTGGAGACACAGGCTGGGAAAGTTTGGAGGCCAGAACCGAAAATCTGAGGAGCAGCTGGAG	2610
DB	2581	CTGGAGACACAGGCTGGGAAAGTTTGGAGGCCAGAACCGAAAATCTGAGGAGCAGCTGGAG	2640
QY	2611	AAGATCAGCCACCAAGACCACTGACAAAGATCGGCTGTGGAATCTGGAGACAAAGATTG	2670
DB	2641	AAGATCAGCCACCAAGACCACTGACAAAGATCGGCTGTGGAATCTGGAGACAAAGATTG	2700
QY	2671	CGGAGGTCACTCTAGACACAGGAGCAGAAATCTGGAGTCAAGGCCAGCTCACAGAG	2730
DB	2701	CGGAGGTCACTCTAGACACAGGAGCAGAAATCTGGAGTCAAGGCCAGCTCACAGAG	2760
QY	2731	CTACAGCTTCCCTCGAGGAGCGGAGTGCACGTTGACAGCCCTGCAAGGCTGCACGGCG	2790
DB	2761	CTACAGCTTCCCTCGAGGAGCGGAGTGCACGTTGACAGCCCTGCAAGGCTGCACGGCG	2820
QY	2791	GCCCTGGAGAGCCAGCTTCGCCAGCGGAAAGACAGAGCTGGAAGAGACCAAGCAAGACT	2850
DB	2821	GCCCTGGAGAGCCAGCTTCGCCAGCGGAAAGACAGAGCTGGAAGAGACCAAGCAAGACT	2880
QY	2851	GAAGGAGATCCAGGCACCTCAGGCACATAGAGATCAATCCAGGCCAAATTTGATGCT	2910
DB	2881	GAAGGAGATCCAGGCACCTCAGGCACATAGAGATCAATCCAGGCCAAATTTGATGCT	2940
QY	2911	CTTCGTAAACAGCTACTGTAAATCACAGACCTGGAGAGCAGCTAAACCCAGCTGACCGAG	2970
DB	2941	CTTCGTAAACAGCTACTGTAAATCACAGACCTGGAGAGCAGCTAAACCCAGCTGACCGAG	3000
QY	2971	GACACGCTGAACCTCAACCAAAAATTTCTACTGTGCCAAAACAACTCGATGAGGCTTCT	3030
DB	3001	GACACGCTGAACCTCAACCAAAAATTTCTACTGTGCCAAAACAACTCGATGAGGCTTCT	3060
QY	3031	GGCGCAACGACGAGATTGTACACTCGCAAGTGAAGTGAACCATCTCCCGCGGAGATC	3090
DB	3061	GGCGCAACGACGAGATTGTACACTCGCAAGTGAAGTGAACCATCTCCCGCGGAGATC	3120
QY	3091	ACGGAAACGAGAGATCGAGCTTACCAGCCAGAGCAACAGATGAGGCTCTGAAGACCCAG	3150
DB	3121	ACGGAAACGAGAGATCGAGCTTACCAGCCAGAGCAACAGATGAGGCTCTGAAGACCCAG	3180
QY	3151	TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA	3210
DB	3181	TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA	3240

QY 3211 AAAGAGCGGAGTGGAGGCGCTCGAGAGCGTCTCGGTGATGAGAAATCCAGTTTGAG 3270
DB 3241 AAAGAGCGGAGTGGAGGCGCTCGAGAGCGTCTCGGTGATGAGAAATCCAGTTTGAG 3300
QY 3271 TGTGGGTTCCGAGAGTGCAGAGAAATGCTGGACACCGAGAGAAACAGAGCGAGGCGAGAGCC 3330
DB 3301 TGTGGGTTCCGAGAGTGCAGAGAAATGCTGGACACCGAGAGAAACAGAGCGAGGCGAGAGCC 3360
QY 3331 GATCAGGGGATCAGGAGTCTCGGAGTGGTGGAGTGGCAGTGAAGGAGCAGAGGCT 3390
DB 3361 GATCAGGGGATCAGGAGTCTCGGAGTGGTGGAGTGGCAGTGAAGGAGCAGAGGCT 3420
QY 3391 GAGATTCTCGTCTGACAGAGGCTCTCAAGAGCAGAGTGAAGGCGGAGAGCCTCTCT 3450
DB 3421 GAGATTCTCGTCTGACAGAGGCTCTCAAGAGCAGAGTGAAGGCGGAGAGCCTCTCT 3480
QY 3451 GACAGCTCAATGACCTGGAGAGAGATGCTATGCTTGAATGAATGATGCCGAGCTTA 3510
DB 3481 GACAGCTCAATGACCTGGAGAGAGATGCTATGCTTGAATGAATGATGCCGAGCTTA 3540
QY 3511 CAGCAGAGCTGAGAGTGAACAGAGAGTCAACAGAGGCTTCTGGAAGAGCAGAGCAA 3570
DB 3541 CAGCAGAGCTGAGAGTGAACAGAGAGTCAACAGAGGCTTCTGGAAGAGCAGAGCAA 3600
QY 3571 TTACAGCAGAGATGAGACTGCGAGAAATCAATTTTCGTCTGACTCAAGAGCTGCAA 3630
DB 3601 TTACAGCAGAGATGAGACTGCGAGAAATCAATTTTCGTCTGACTCAAGAGCTGCAA 3660
QY 3631 GAGACTCTAGATCGGCTGATCTACTGAACAGAGAGTGAAGTGAAGTATCAGCTG 3690
DB 3661 GAGACTCTAGATCGGCTGATCTACTGAACAGAGAGTGAAGTGAAGTATCAGCTG 3720
QY 3691 GAAACATTCAGGTTCTTATCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3750
DB 3721 GAAACATTCAGGTTCTTATCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3780
QY 3751 CAACCAACTCATTGATTTCTGCAAGCCAAATGGAACCACTCTTAAAGAGAAAG 3810
DB 3781 CAACCAACTCATTGATTTCTGCAAGCCAAATGGAACCACTCTTAAAGAGAAAG -- 3838
QY 3811 GGTTATTATTAGTCGAGGAAAGAGGACCTGCTTTTACCACACAGAGTTCCTCTGCAGTAC 3870
DB 3839 -----AGTTCCTCTGCAGTAC 3855
QY 3871 AATGAGCTGAAGTGGCCCTGGAGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3930
DB 3856 AATGAGCTGAAGTGGCCCTGGAGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3915
QY 3931 CTTTCAGAGACCCGATCGAGCTCCGGTCCGCCGGGAGGAGAGTCCCAACCGCAAGCA 3990
DB 3916 CTTTCAGAGACCCGATCGAGCTCCGGTCCGCCGGGAGGAGAGTCCCAACCGCAAGCA 3975
QY 3991 ACAGAACCCACACCGATCCACCGAGCCACCGCGAGGAGCAGATCGCCATGTCGCGC 4050
DB 3976 ACAGAACCCACACCGATCCACCGAGCCACCGCGAGGAGCAGATCGCCATGTCGCGC 4035
QY 4051 ATCGTGGGTGCGCAGAGCAGACCCAGTGGCCATCAGAGCTGTGCCCCCGCCATCCAGC 4110
DB 4036 ATCGTGGGTGCGCAGAGCAGACCCAGTGGCCATCAGAGCTGTGCCCCCGCCATCCAGC 4095
QY 4111 CGCAGAAAGAGTCTTCAATCCAGAGGAAATTTAGTCGCGCTTTAAGGAACCGCATGCAAC 4170
DB 4096 CGCAGAAAGAGTCTTCAATCCAGAGGAAATTTAGTCGCGCTTTAAGGAACCGCATGCAAC 4155
QY 4171 CACAATATTCCTCACCGATTCAAGTGAAGTGAATGAGAGTGAAGTGAAGTGAAGTGAAG 4230
DB 4156 CACAATATTCCTCACCGATTCAAGTGAAGTGAATGAGAGTGAAGTGAAGTGAAGTGAAG 4215
QY 4231 TGTCTGGATACCGTGTGACCTTGGAGCCAGGATCCAAATGCTTGAATGTCAGGTGATG 4290
DB 4216 TGTCTGGATACCGTGTGACCTTGGAGCCAGGATCCAAATGCTTGAATGTCAGGTGATG 4275
QY 4291 TGTCAACCCCAAGTGTCTCCAGCTGCTGCCAGCCACCTGCGGCTTGCTGCTGAAATGCC 4350

DB 4276 TGTCAACCCCAAGTGTCTCCAGTGTCTGCCAGCACCTGCGCTTGCTGTGATATGCC 4335
QY 4351 ACACACTTCAACGAGGCTTCTGCCGTGACAAATGAATGAATCTCCAGGTCTCCAGACCAAG 4410
DB 4336 ACACACTTCAACGAGGCTTCTGCCGTGACAAATGAATGAATCTCCAGGTCTCCAGACCAAG 4395
QY 4411 GAGCCACAGCAGCTTGCACCTGCAAGGTTGGATGAAGTGCACAGGAATAACAAACGA 4470
DB 4396 GAGCCACAGCAGCTTGCACCTGCAAGGTTGGATGAAGTGCACAGGAATAACAAACGA 4455
QY 4471 GGAAGAGAGCTGGGACAGGAAGTACATTTCTGAGGGATCAAAAGTCTCATTTAT 4530
DB 4456 GGAAGAGAGCTGGGACAGGAAGTACATTTCTGAGGGATCAAAAGTCTCATTTAT 4515
QY 4531 GACATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4590
DB 4516 GACATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4575
QY 4591 GACGGGATGATCTATTTCTATGTTGGTCTTCCGAACTTCGAAATACAGCCAAA 4650
DB 4576 GACGGGATGATCTATTTCTATGTTGGTCTTCCGAACTTCGAAATACAGCCAAA 4635
QY 4651 GCA----- 4653
DB 4636 GCAGATGTCANTACATCTGAAGATGGAATCTCACCCGACACCACTGCTGCGCCGGG 4695
QY 4654 ----- 4653
DB 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACGCTTA 4755
QY 4654 -----GAAAAACAGAGCTGATGCTAACTG 4680
DB 4756 GAATCAGTTGTCAGGTGGAGAGTTTCTAGGAAAAAGCAGAGTGTGCTAACTG 4815
QY 4681 CTTGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACTGCACGCTG 4740
DB 4816 CTTGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACTGCACGCTG 4875
QY 4741 CCCTTCAGTACCAAGTGTGTTGGTGGGCAACGAGGAGGGCTCTACGCCCTGGAATGTC 4800
DB 4876 CCCTTCAGTACCAAGTGTGTTGGTGGGCAACGAGGAGGGCTCTACGCCCTGGAATGTC 4935
QY 4801 TTGAAACTCCCTTAACCCATGTCAGAGAAATGGAGCAGTCTTCCAAATTTATATATC 4860
DB 4936 TTGAAACTCCCTTAACCCATGTCAGAGAAATGGAGCAGTCTTCCAAATTTATATATC 4995
QY 4861 AAGGACCTGGAGAGTACTCATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGAC 4920
DB 4996 AAGGACCTGGAGAGTACTCATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGAC 5055
QY 4921 GTGAGAGAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGCACTCTCA 4980
DB 5056 GTGAGAGAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGCACTCTCA 5115
QY 4981 CCCAACATTTTGAAGCTGTCAAGGCTGCACTTGTTTGGGGCAGGCAAGATTGGAAC 5040
DB 5116 CCCAACATTTTGAAGCTGTCAAGGCTGCACTTGTTTGGGGCAGGCAAGATTGGAAC 5175
QY 5041 GGGCTGTGATCTGTGAGCCATGCCAGCAAGTGTGCTATCTCCGCTACACAGAAAC 5100
DB 5176 GGGCTGTGATCTGTGAGCCATGCCAGCAAGTGTGCTATCTCCGCTACACAGAAAC 5235
QY 5101 CTCAGCAAAATCTCATCTCCGAAAGAGATAGAGACTTCAGAGCCCTGCACTGTATCCAC 5160
DB 5236 CTCAGCAAAATCTCATCTCCGAAAGAGATAGAGACTTCAGAGCCCTGCACTGTATCCAC 5295
QY 5161 TTCAACCAATTAAGTATCTCATTTGGAACCAATAAATTTACGAAATCGACATGAAGCAG 5220
DB 5296 TTCAACCAATTAAGTATCTCATTTGGAACCAATAAATTTACGAAATCGACATGAAGCAG 5355
QY 5221 TACAGCTCGAGGAAATCTCGGATGAAGTGAACATTTCTTGGCAGCTGTGTTTGGC 5280

Db 5356 TACACGCTCGAGGAATTCCTGGATAGAATGACCAATTCCTTGGCACCCTGCTGTGTTGCC 5415
QY 5281 GCCTCTTCCAAACAGCTTCCCTGCTCTCAATCGTGCAGTCAACAGCGCAGCGGAGAGAG 5340
Db 5416 GCCTCTTCCAAACAGCTTCCCTGCTCTCAATCGTGCAGTCAACAGCGCAGCGGAGAGAG 5475
QY 5341 GAGTACTGCTGTGTTTCCAGAAATTTGAGAGTTCCTGCTGATTTCTTACGGAAGAGCTAGC 5400
Db 5476 GAGTACTGCTGTGTTTCCAGAAATTTGAGAGTTCCTGCTGATTTCTTACGGAAGAGCTAGC 5535
QY 5401 CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTGGCCCTTTGGCTTACAGAGAAACCTAT 5460
Db 5536 CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTGGCCCTTTGGCTTACAGAGAAACCTAT 5595
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCA 5520
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCA 5655
QY 5521 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCCGCTACCTGGGCCCTGCCATT 5580
Db 5656 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCCGCTACCTGGGCCCTGCCATT 5715
QY 5581 TCCTCAGGAGCGATTTACTTTGGCGTCTCATACAGGATAAATTAGGGGTCAATTTGCTGTC 5640
Db 5716 TCCTCAGGAGCGATTTACTTTGGCGTCTCATACAGGATAAATTAGGGGTCAATTTGCTGTC 5775
QY 5641 AAGGGAACCTGCTGAAGAGTCCGGCACTGAACACCCGGGGCCGCTCCACCTCCCGC 5700
Db 5776 AAGGGAACCTGCTGAAGAGTCCGGCACTGAACACCCGGGGCCGCTCCACCTCCCGC 5835
QY 5701 AGCAGCCCCCAAGAGGAGCGCCACCCAGCTACACGACATCACCAAGCGCTGGCC 5760
Db 5836 AGCAGCCCCCAAGAGGAGCGCCACCCAGCTACACGACATCACCAAGCGCTGGCC 5895
QY 5761 TCAGCCCGAGCGCGCCGAGGCGCCAGCCAGCCCGGAGAGCCCAAGCACACCCACCGC 5820
Db 5896 TCAGCCCGAGCGCGCGCCGAGGCGCCAGCCAGCCCGGAGAGCCCAAGCACACCCACCGC 5955
QY 5821 TACCGCGGGCGGAGCGAGTCCGAGGAGCAAGTCTCTGGCGGCCCTGAGGCGA 5880
Db 5956 TACCGCGGGCGGAGCGAGTCCGAGGAGCAAGTCTCTGGCGGCCCTGAGGCGA 6015
QY 5881 GAGAACTCCCCCGGCGGATGCTCAGCAGCGGAGAGCGGTCCCGGAGGCTGTTT 5940
Db 6016 GAGAACTCCCCCGGCGGATGCTCAGCAGCGGAGAGCGGTCCCGGAGGCTGTTT 6075
QY 5941 GAAGACAGCAGAGGGGCGCGTCCCTGCGGAGCGGTGAGGACCCCGCTGTCAGGTTG 6000
Db 6076 GAAGACAGCAGAGGGGCGCGTCCCTGCGGAGCGGTGAGGACCCCGCTGTCAGGTTG 6135
QY 6001 AACAGGGGAGAGGGCA 6017
Db 6136 AACAGGTGAGGAGCA 6152

RESULT 5
AAL55214
ID AAL55214 standard; DNA; 6165 BP.
XX AC AAL55214;
XX DT 01-MAY-2003 (first entry)
XX DE Human CR1K encoding DNA sequence, SEQ ID No 1.
XX KW Anorectic; hypotensive; cardiatic; antilipaeamic; cerebroprotective;
KW antitgout; osteopathic; antiarthritic; cyostatic; antidepressant;
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;
KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;
KW central nervous system disorder; chronic obstructive pulmonary disease;
KW diabetes; pain; gene; ds.

XX Homo sapiens.
OS XX Key Location/Qualifiers
FH CDS 1..6165
FT /*tag= a
FT /product= "Human CR1K protein"
XX PN WO2003004523-A1.
XX PD 16-JAN-2003.
XX PF 28-JUN-2002; 2002WO-EP007156.
XX PR 02-JUL-2001; 2001US-0301841P.
XX PR 11-DEC-2001; 2001US-0338651P.
XX PR 25-APR-2002; 2002US-0375014P.
XX PA (FARB) BAYER AG.
XX PI Zhu Z;
XX WPI; 2003-221576/21.
XX P-PSDB; AAO26959.
XX New human citron rho/rac-interacting kinase (CR1K) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX Example 1; Fig 1; 237pp; English.
CC The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CR1K)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CR1K dysfunction such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
CC cancer including endometrial, breast, prostate and colon cancer),
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC disease), chronic obstructive pulmonary disease, or diabetes. These can
CC also be used to treat pain associated with the disorders. The human CR1K
CC polypeptide is also useful in diagnostic assays or in genetic testing.
CC The expression vector or the reagent is useful in preparing a medicament
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a
CC central nervous system disorder, or chronic obstructive pulmonary
CC disease. The fusion protein is useful for generating antibodies against a
CC CR1K polypeptide and for use in various assay systems. The methods are
CC useful in producing and detecting the polynucleotide and polypeptide and
CC in screening for agents that modulate the activity of the human CR1K
CC polypeptide. This polynucleotide sequence represents a DNA sequence
CC encoding a human CR1K protein of the invention
XX
SQ Sequence 6165 BP; 1735 A; 1549 C; 1680 G; 1201 T; 0 U; 0 Other;

Query Match 86.2%; Score 5666.2; DB 8; Length 6165;
Best Local Similarity 95.9%; Pred.No. 0;
Matches 5947; Conservative 0; Mismatches 13; Indels 243; Gaps 3;
QY 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAACCCATT 78
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAACCCATT 60
QY 79 GCCAGCGGGCTCCAGGCTGAATCTTTCTTCAGGGGAAACACCCCTTATGACTCAA 138
Db 61 GCCAGCGGGCTCCAGGCTGAATCTTTCTTCAGGGGAAACACCCCTTATGACTCAA 120
QY 139 CAGCAGATGCTCCTCTTTCCGAGAGGGATATTAGATCCCTCTTTGCTCTTTGAA 198
Db 121 CAGCAGATGCTCCTCTTTCCGAGAGGGATATTAGATCCCTCTTTGCTCTTTGAA 180

Qy	199	GAATG	CAGT	CAGC	CTGCT	CTGAT	GAA	GATT	TAAG	CAC	GTG	AG	CAACT	TTGT	CCG	GAAG	TAT	258	
Db	181	GAATG	CAGT	CAGC	CTGCT	CTGAT	GAA	GATT	TAAG	CAC	GTG	AG	CAACT	TTGT	CCG	GAAG	TAT	240	
Qy	259	TCC	GAC	ACCA	TAG	CTG	AGT	TAC	AGG	AGCT	CCAG	CCCT	CGCG	AAAG	ACTTT	CG	AGTC	AGA	318
Db	241	TCC	GAC	ACCA	TAG	CTG	AGT	TAC	AGG	AGCT	CCAG	CCCT	CGCG	AAAG	ACTTT	CG	AGTC	AGA	300
Qy	319	AGT	CTT	G	TAG	TTG	TG	TC	AGT	TC	AGT	TC	AGT	TC	AGT	TC	AGT	TC	378
Db	301	AGT	CTT	G	TAG	TTG	TG	TC	AGT	TC	AGT	TC	AGT	TC	AGT	TC	AGT	TC	360
Qy	379	GAC	ATC	TAT	GC	TAT	CA	AGT	GAT	CA	AG	AA	AGC	TTT	AT	TG	CC	CC	438
Db	361	GAC	ATC	TAT	GC	TAT	CA	AGT	GAT	CA	AG	AA	AGC	TTT	AT	TG	CC	CC	420
Qy	439	TTTTT	T	GAG	GA	GCG	GA	CA	TAT	TAT	CTG	AA	GCA	CA	AG	CCCG	TG	GAAT	498
Db	421	TTTTT	T	GAG	GA	GCG	GA	CA	TAT	TAT	CTG	AA	GCA	CA	AG	CCCG	TG	GAAT	480
Qy	499	CAG	TAT	G	CC	TTT	C	AGG	CA	AAAT	CA	CTT	TAT	CTG	AT	G	AG	GA	558
Db	481	CAG	TAT	G	CC	TTT	C	AGG	CA	AAAT	CA	CTT	TAT	CTG	AT	G	AG	GA	540
Qy	559	GAC	TTG	CT	G	CAC	TTT	T	GA	TAG	AT	G	AG	CA	AGT	TAG	AT	GA	618
Db	541	GAC	TTG	CT	G	CAC	TTT	T	GA	TAG	AT	G	AG	CA	AGT	TAG	AT	GA	600
Qy	619	TAC	CT	AG	CT	G	AT	G	AT	T	T	G	CT	TT	C	A	G	CT	678
Db	601	TAC	CT	AG	CT	G	AT	T	T	G	CT	TT	C	A	G	CT	TT	C	660
Qy	679	GAC	AT	CA	AG	CC	T	G	AA	CA	T	T	C	G	T	T	G	A	738
Db	661	GAC	AT	CA	AG	CC	T	G	AA	CA	T	T	C	G	T	T	G	A	720
Qy	739	GG	AT	CT	G	CC	G	GA	AA	T	G	AA	T	G	CT	GA	T	G	798
Db	721	GG	AT	CT	G	CC	G	GA	AA	T	G	AA	T	G	CT	GA	T	G	780
Qy	799	CC	AG	AT	T	AC	AT	G	CT	CT	G	AA	T	G	CT	CT	G	AA	858
Db	781	CC	AG	AT	T	AC	AT	G	CT	CT	G	AA	T	G	CT	CT	G	AA	840
Qy	859	GG	CT	G	CA	CT	G	AT	G	CT	G	GG	CG	T	GA	T	G	CT	918
Db	841	GG	CT	G	CA	CT	G	AT	G	CT	G	GG	CG	T	GA	T	G	CT	900
Qy	919	TCC	CC	CT	TC	G	AG	GA	AA	CT	CT	TC	G	AG	GA	AA	CT	TC	978
Db	901	TCC	CC	CT	TC	G	AG	GA	AA	CT	CT	TC	G	AG	GA	AA	CT	TC	960
Qy	979	TTTTT	G	AA	AT	T	C	AG	AT	G	CT	TC	GA	T	G	CT	TC	GA	1038
Db	961	TTTTT	G	AA	AT	T	C	AG	AT	G	CT	TC	GA	T	G	CT	TC	GA	1020
Qy	1039	TTG	TT	G	T	G	CG	CC	CA	GA	AG	GA	GA	CT	GA	AG	GA	AG	1098
Db	1021	TTG	TT	G	T	G	CG	CC	CA	GA	AG	GA	GA	CT	GA	AG	GA	AG	1080
Qy	1099	TCT	AAA	AT	T	G	AT	G	CT	TC	GA	AA	CT	TC	GA	AA	CT	TC	1158
Db	1081	TCT	AAA	AT	T	G	AT	G	CT	TC	GA	AA	CT	TC	GA	AA	CT	TC	1140

Qy	1279	TACAGCAAGGCATCTGGGATTCCTGTGTAGATCTGAGTCTCTTTGTTCGGGTCTGGACTCC	1333
Db	1261	TACAGCAAGGCATCTGGGATTCCTGTGTAGATCTGAGTCTCTTTGTTCGGGTCTGGACTCC	1320
Qy	1339	CTGTCCAAAGACTAGCTCCATCGGAAAGAAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1398
Db	1321	CTGTCCAAAGACTAGCTCCATCGGAAAGAAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1380
Qy	1399	TCTCAGACAAGTGTCAACAGATGGAGCAGGAAATCAACCCGGTTACATCGGAGAGTGTCA	1458
Db	1381	TCTCAGACAAGTGTCAACAGATGGAGCAGGAAATCAACCCGGTTACATCGGAGAGTGTCA	1440
Qy	1459	GAGGTGGAGCTGTCTTAGTCTAGAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1518
Db	1441	GAGGTGGAGCTGTCTTAGTCTAGAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1500
Qy	1519	TCCTCTCTGGAGCAGGACCTTGCTACTCATCATCAGAAATGCAGTAGCTTAAAGCGGAAGT	1578
Db	1501	TCCTCTCTGGAGCAGGACCTTGCTACTCATCATCAGAAATGCAGTAGCTTAAAGCGGAAGT	1560
Qy	1579	TTGGAGCAAGCAGGATGAGGTGTCCGAGGAGGATGACAAAGCACTGCAGCTTCTCCAT	1638
Db	1561	TTGGAGCAAGCAGGATGAGGTGTCCGAGGAGGATGACAAAGCACTGCAGCTTCTCCAT	1620
Qy	1639	GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA	1698
Db	1621	GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA	1680
Qy	1699	GTGGAAGAAATGAGTTGATGATGAATCAAGTTGGAGAAGATCTTCTCAGCAAGAGA	1758
Db	1681	GTGGAAGAAATGAGTTGATGATGAATCAAGTTGGAGAAGATCTTCTCAGCAAGAGA	1740
Qy	1759	CGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTGCTGCTGAAGAAATCAAG	1818
Db	1741	CGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTGCTGCTGAAGAAATCAAG	1800
Qy	1819	CGGAAAGCGACGAATGTCAGCAATAACTGTTGAAGGCTAAGGATCAAGGGAAGCTGAA	1878
Db	1801	CGGAAAGCGACGAATGTCAGCAATAACTGTTGAAGGCTAAGGATCAAGGGAAGCTGAA	1860
Qy	1879	GTGGGAGATATCGGAAACTGGAGAAGATCAATGCTTGAGCAGCAGCTCAAAATTCAGGAG	1938
Db	1861	GTGGGAGATATCGGAAACTGGAGAAGATCAATGCTTGAGCAGCAGCTCAAAATTCAGGAG	1920
Qy	1939	CTCAAGAGAAACTGGAGAAAGCT-----	1962
Db	1921	CTCAAGAGAAACTGGAGAAAGCTTAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAG	1980
Qy	1963	-----GCAAGAGGACGAGCCGAGAGGAGCTGGAGAAGCTGCAGAAACCGAGAG	2010
Db	1981	AATATCGCCAGGCAAAAGGACGAGCCGAGAGGAGCTGGAGAAGCTGCAGAAACCGAGAG	2040
Qy	2011	GATTCCTCTGAAGGCATCAGAAAGAGCTGGTGGAAAGCTTGAGGAAACGCCCAATTCCTG	2070
Db	2041	GATTCCTCTGAAGGCATCAGAAAGAGCTGGTGGAAAGCTTGAGGAAACGCCCAATTCCTG	2100
Qy	2071	GAGAACAAAGTAAAGAGACTAGAGACCATGGACGCTAGAGAAAACAGACTGAAGGATGAC	2130
Db	2101	GAGAACAAAGTAAAGAGACTAGAGACCATGGACGCTAGAGAAAACAGACTGAAGGATGAC	2160
Qy	2131	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTCGGAGCTCGAGAG	2190
Db	2161	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTCGGAGCTCGAGAG	2220
Qy	2191	AAACATCGGAGGCCCAAGTCTCAGGCCACGACCTAGAAAGTGCACCTGAAAACAGAAAGAG	2250
Db	2221	AAACATCGGAGGCCCAAGTCTCAGGCCACGACCTAGAAAGTGCACCTGAAAACAGAAAGAG	2280
Qy	2251	CAGACATATCAGGAAAAGATTAAAGTTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2310
Db	2281	CAGACATATCAGGAAAAGATTAAAGTTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2340
Qy	2311	AAGGAGACCTGCGAAGACATGATGCGAGAGACGAGGAGGAGGCCCATGAGAAAGGCCAAA	2370

Db 2341 AAGGAGACACTGGAGAACATGATGTCAGAGACACGAGGAGGAGGCCCATGAGAAAGGCGCAA 2400
QY 2371 ATTCTCAGCGAAACAGAGCGCATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2430
Db 2401 ATTCTCAGCGAAACAGAGCGCATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2460
QY 2431 CAGAGATTGTGGAACTGCTGAAGCCATAAATCTGAGCAAAATAGCAGTCTTTTACC 2490
Db 2461 CAGAGATTGTGGAACTGCTGAAGCCATAAATCTGAGCAAAATAGCAGTCTTTTACC 2520
QY 2491 CAAAGAAACATGAAGCCCCAAGAAGAGATGATTTCTGAACTCAGGCAACAGAAAAATTTTAC 2550
Db 2521 CAAAGAAACATGAAGCCCCAAGAAGAGATGATTTCTGAACTCAGGCAACAGAAAAATTTTAC 2580
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGGCCGAGACCGGAAAACTGGAGGAGCAGCTGGAG 2610
Db 2581 CTGGAGACACAGCTGGGAAGTTGGAGGCCGAGAACCGGAAAACTGGAGGAGCAGCTGGAG 2640
QY 2611 AAGATCAGCCACCAAGACACACAGTGAAGAATCGCTCTGAACTGGAACTGGAGCAAGATTG 2670
Db 2641 AAGATCAGCCACCAAGACACACAGTGAAGAATCGCTCTGAACTGGAGCAAGATTG 2700
QY 2671 CGGAGGTCAGTCTAGAGCACGAGGACGAGAACTGGAGCTCAAGCGCCAGCTCACAG 2730
Db 2701 CGGAGGTCAGTCTAGAGCACGAGGACGAGAACTGGAGCTCAAGCGCCAGCTCACAG 2760
QY 2731 CTACAGCTCTCCCTGCAGAGCGCGAGTCACAGTTGACAGCCCTCAGGCTGCACGGGG 2790
Db 2761 CTACAGCTCTCCCTGCAGAGCGCGAGTCACAGTTGACAGCCCTCAGGCTGCACGGGG 2820
QY 2791 GCCCTGGAGAGCCAGCTTCGCCAGGCGGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2850
Db 2821 GCCCTGGAGAGCCAGCTTCGCCAGGCGGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2880
QY 2851 GAAGAGGATCCAGCACCTCAGGCACATAGAGATGAATCCAGGCGAAATTTGATGCT 2910
Db 2881 GAAGAGGATCCAGGCATCAGGCACATAGAGATGAATCCAGGCGAAATTTGATGCT 2940
QY 2911 CTTCTGAACAGCTGTACTGTATCAAGACCTGGAGGAGCAGTAAACAGCTGACCCGAG 2970
Db 2941 CTTCTGAACAGCTGTACTGTATCAAGACCTGGAGGAGCAGTAAACAGCTGACCCGAG 3000
QY 2971 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030
Db 3001 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060
QY 3031 GCGGCCAACGACGAGATTGTACACTGCGAAGTGAAGTGGACCATCTCCGCGGGAGATC 3090
Db 3061 GCGGCCAACGACGAGATTGTACACTGCGAAGTGAAGTGGACCATCTCCGCGGGAGATC 3120
QY 3091 ACGGAACGAGAGATGACGCTTACCGCCAGAGCAAAACGATGGAGCTCTGAAGACCCAG 3150
Db 3121 ACGGAACGAGAGATGACGCTTACCGCCAGAGCAAAACGATGGAGCTCTGAAGACCCAG 3180
QY 3151 TGCACCATGCTGGAGAAACAGTATGGAATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210
Db 3181 TGCACCATGCTGGAGAAACAGTATGGAATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240
QY 3211 AAAGACGCGCAGTGGAGGCGCTGGAGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGA 3270
Db 3241 AAAGACGCGCAGTGGAGGCGCTGGAGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGA 3300
QY 3271 TGTGCGGTTTCGAGAGTTCAGAGAAATGCTGGACACCGAGAAAACAGAGCAGGCGAGAGCC 3330
Db 3301 TGTGCGGTTTCGAGAGTTCAGAGAAATGCTGGACACCGAGAAAACAGAGCAGGCGAGAGCC 3360
QY 3331 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAGGCT 3390
Db 3361 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAGGCT 3420
QY 3391 GAGATTCTCGCTCTGACAGGCTCTCAAGAGCAGAACTGAAAGCCGAGAGCCTCTCT 3450

Db 3421 GAGATTCTCGCTCTGACAGCGGCTCTCAAGAGCAGAGAGCTGAAGCCGAGAGCCTCTCT 3480
QY 3451 GACAAGCTCAATGACTCTGGAGAAAGACATGCTATGTTGAAATGAATGCCCGAAGCTTA 3510
Db 3481 GACAAGCTCAATGACTCTGGAGAAAGACATGCTATGTTGAAATGAATGCCCGAAGCTTA 3540
QY 3511 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAGACCCAAA 3570
Db 3541 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAGACCCAAA 3600
QY 3571 TTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTCAA 3630
Db 3601 TTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTCAA 3660
QY 3631 GAACTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGGAAGGACATATTCTCAA 3690
Db 3661 GAACTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGGAAGGACATATTCTCAA 3720
QY 3691 GAAACATTTCAAGTTCTCTATTCTATGAAAGGTGAAAAATGGAAGGACATATTCTCAA 3750
Db 3721 GAAACATTTCAAGTTCTCTATTCTCATGAAAGGTGAAAAATGGAAGGACATATTCTCAA 3780
QY 3751 CAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGGAACCAACCTGCTAAAAAGAAAA 3810
Db 3781 CAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGGAACCAACCTGCTAAAAAGAAAA -- 3838
QY 3811 GGTTTATTAGTCGACGAAAAAGAGGCCCTGCTTTTACCCACACAGGTTCTCTCAGTAC 3870
Db 3839 -----AGTTCTCTCAGTAC 3855
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTCGTGTGCAGAGCTAGAGAAAGCC 3930
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTCGTGTGCAGAGCTAGAGAAAGCC 3915
QY 3931 CTTCAGAGACCCGATCGAGCTCCGTCGCGCGGAGGAAGCTGCCACCGCAAGCA 3990
Db 3916 CTTCAGAGACCCGATCGAGCTCCGTCGCGCGGAGGAAGCTGCCACCGCAAGCA 3975
QY 3991 ACGACCAACCCACACCCATCCACCCAGCCACCGCGAGGAGCAGATCGGCATGTCGCC 4050
Db 3976 ACGACCAACCCACACCCATCCACCCAGCCACCGCGAGGAGCAGATCGGCATGTCGCC 4035
QY 4051 ATCGTGGTGCAGAGACACCGCCAGTGCATGAGCCTGCTGGCGCCGCCATCCAGC 4110
Db 4036 ATCGTGGTGCAGAGACACCGCCAGTGCATGAGCCTGCTGGCGCCGCCATCCAGC 4095
QY 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGTCTTAAAGGAACGATGCAC 4170
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGTCTTAAAGGAACGATGCAC 4155
QY 4171 CACAATATTCCTCAACCAATCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4230
Db 4156 CACAATATTCCTCAACCAATCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4215
QY 4231 TGTCTGGATACCGTGCATTTGGACGCCAGGCATCCAAATGCTCTCGAATGTCAGGTGATG 4290
Db 4216 TGTCTGGATACCGTGCATTTGGACGCCAGGCATCCAAATGCTCTCGAATGTCAGGTGATG 4275
QY 4291 TGTCAACCCCAAGTCTCCAGTGTCTGCCAGCCACCTGCGGCTTGCCTGTGATATGCC 4350
Db 4276 TGTCAACCCCAAGTCTCCAGTGTCTGCCAGCCACCTGCGGCTTGCCTGTGATATGCC 4335
QY 4351 ACACATTTACCGAGGCTTCTGCCGTGACAAAATGAACTCCCGAGGTCTCCAGACCAAG 4410
Db 4336 ACACATTTACCGAGGCTTCTGCCGTGACAAAATGAACTCCCGAGGTCTCCAGACCAAG 4395
QY 4411 GAGCCACGACGACGCTTGCACCTGGAAGGTGGAAGTGCCTCCAGGAATTAACAAACGA 4470
Db 4396 GAGCCACGACGACGCTTGCACCTGGAAGGTGGAAGTGCCTCCAGGAATTAACAAACGA 4455
QY 4471 GGAAGCAAGGCTGGGACAGAAATCAATGCTCTGGAGGGAATCAAAAGTCTCTCATTTAT 4530
Db 4456 GGAAGCAAGGCTGGGACAGAAATCAATGCTCTGGAGGGAATCAAAAGTCTCTCATTTAT 4515

```
QY 4531 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4590
Db 4516 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4575
QY 4591 GACGGGATGTATCTATTTCATGTGTGCGTGTGGTCTTCCGAATCGCAATACAGCAAA 4650
Db 4576 GACGGGATGTATCTATTTCATGTGTGCGTGTGGTCTTCCGAATCGCAATACAGCAAA 4635
QY 4651 GCA----- 4653
Db 4636 GCAGATGCCATACATACACTGAAGATGGAATCTCACCCGCACACCACTGCTGCGCCCGG 4695
QY 4654 ----- 4653
Db 4696 AGAACCCCTACTTGTAGTCCAGCTCCCTGACAAAACAGCGCTGGGTCAACCGCTTA 4755
QY 4654 -----GAAAACGAGAGCTGATGCTAACTG 4680
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAACTG 4815
QY 4681 CTTGGAAATCCCTGCTGAAACTGGAGGTGATGACCGTCTAGACATGAACCTGACGCTG 4740
Db 4816 CTTGGAAATCCCTGCTGAAACTGGAGGTGATGACCGTCTAGACATGAACCTGACGCTG 4875
QY 4741 CCCTTCAGTGAACCAAGTGTGTTGGTGGCCACCGAGGAAAGGCTCTACGCCCTGAATGTC 4800
Db 4876 CCCTTCAGTGACAGAGTGTGTTGGTGGCCACCGAGGAAAGGCTCTACGCCCTGAATGTC 4935
QY 4801 TTGAAAATCCCTTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATTTC 4860
Db 4936 TTGAAAATCCCTTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATTTC 4995
QY 4861 AAGGACCTGGAGAGCTACTCATATAGCAGGAGAAAGAGCGGCACTGTCTTGTGGAC 4920
Db 4996 AAGGACCTGGAGAGCTACTCATATAGCAGGAGAAAGAGCGGCACTGTCTTGTGGAC 5055
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4980
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115
QY 4981 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGAGGCAAGATTGAGAAC 5040
Db 5116 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGAGGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGCACTGTGAGCCATGCCACGAAAGTCTGTCATTTCTCGCTACAAAGAAAC 5100
Db 5176 GGGCTCTGCACTGTGAGCCATGCCACGAAAGTCTGTCATTTCTCGCTACAAAGAAAC 5235
QY 5101 CTCAGAAATPACTGCTATCGGAAGAGATAGAGACCTCAGAGCCCTGACGCTGTATCCAC 5160
Db 5236 CTCAGAAATPACTGCTATCGGAAGAGATAGAGACCTCAGAGCCCTGACGCTGTATCCAC 5295
QY 5161 TTCACCAATPACAGTATCCTCATTTGGAACCAATAAAATTTCTAGAAATCGACATGAAGCAG 5220
Db 5296 TTCACCAATPACAGTATCCTCATTTGGAACCAATAAAATTTCTAGAAATCGACATGAAGCAG 5355
QY 5221 TACACGCTCGAGAAATTCCTGGATAGAAATGACCATTCCTTGGCACCTGCTGTGTTGCC 5280
Db 5356 TACACGCTCGAGAAATTCCTGGATAGAAATGACCATTCCTTGGCACCTGCTGTGTTGCC 5415
QY 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGAGCGAGGCGAGAGAG 5340
Db 5416 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGAGCGAGGCGAGAGAG 5475
QY 5341 GAGTACTTGTGTTTCCAGAAATTTGGAGTGTTCGTGGATTCTTACGAAAGAGCTAGC 5400
Db 5476 GAGTACTTGTGTTTCCAGAAATTTGGAGTGTTCGTGGATTCTTACGAAAGAGCTAGC 5535
QY 5401 CGCACAGAGATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGGCCTTACAGAGAACCTAT 5460
Db 5536 CGCACAGAGATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGGCCTTACAGAGAACCTAT 5595
```

```
QY 5461 CTGTTTGTGACCCCACTTCAACTCTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5520
Db 5596 CTGTTTGTGACCCCACTTCAACTCTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5655
QY 5521 GGGACCCCTGCCCAGAGCGTACTTGAGCATCCGAAACCCGGCTACCTTGGGCCCTTGCATT 5580
Db 5656 GGGACCCCTGCCCAGAGCGTACTTGAGCATCCGAAACCCGGCTACCTTGGGCCCTTGCATT 5715
QY 5581 TCCTCAGAGAGGATTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTGCTGC 5640
Db 5716 TCCTCAGAGAGGATTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTGCTGC 5775
QY 5641 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCCGC 5700
Db 5776 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCCGC 5835
QY 5701 AGCAGCCCAACAAAGAGAGGCCCAACCAAGTACAAACAGAGACATCACAAAGCGGTGGCC 5760
Db 5836 AGCAGCCCAACAAAGAGAGGCCCAACCAAGTACAAACAGAGACATCACAAAGCGGTGGCC 5895
QY 5761 TCCAGCCAGCGCCGCCAGAGGCCCAACCAAGGCCCCAGCCAGCCGAGAGCAACACCCACCGC 5820
Db 5896 TCCAGCCAGCGCCGCCAGAGGCCCAACCAAGGCCCCAGCCAGCCGAGAGCAACACCCACCGC 5955
QY 5821 TACCGAGAGGGCGGAGCCGAGCTCGCAGGAGCAAGTCTCTGCGCCGCCCTTGGAGCGA 5880
Db 5956 TACCGAGAGGGCGGAGCCGAGCTCGCAGGAGCAAGTCTCTGCGCCGCCCTTGGAGCGA 6015
QY 5881 GAGAAGTCCCCCGCCCGATGCTCAGCACCGGAGAGAGCGGTCCCGGGAGGCTGTTT 5940
Db 6016 GAGAAGTCCCCCGCCCGATGCTCAGCACCGGAGAGAGCGGTCCCGGGAGGCTGTTT 6075
QY 5941 GAACAGACAGCAGGGGCCGCTGCTGCGGGAGCCCGTGGAGACCCCGCTGTCCCGAGTG 6000
Db 6076 GAACAGACAGCAGGGGCCGCTGCTGCGGGAGCCCGTGGAGACCCCGCTGTCCCGAGTG 6135
QY 6001 AACAGGGAAGAGGCGAGAGTGC 6023
Db 6136 AACAGGTCGGGACCACTTTC 6158

RESULT 6
ABQ78870
ID ABQ78870 standard; cDNA; 6165 BP.
XX AC ABQ78870;
XX DT 10-OCT-2002 (first entry)
XX OS Homo sapiens.
XX DE Human kinase cDNA #1.
XX KW Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;
XX KW Citron rho-interacting kinase; gene therapy; mental disorder; cancer;
XX KW gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..6165
XX FT /*tag= a
XX FT /product= "Kinase"
XX FT replace(5218,G)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT replace(6065,G)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX PN WO200259325-A2.
XX PD 01-AUG-2002.
XX PF 20-DEC-2001; 2001WO-US050497.
```

XX	27-DEC-2000; 2000US-0258335P.
XX	(LEXI-) LEXICON GENETICS INC.
XX	Yu X, Miranda M, Friddle CJ;
PI	WPI; 2002-599796/64.
DR	P-PSDB; ABB81927.
XX	
XX	Novel polynucleotide encoding human proteins that are structurally
PT	similar to animal kinases, useful for drug screening, diagnosis, in gene
PT	therapy of disorders and diseases e.g. cancer and pharmacogenetic
PT	applications.
XX	
PS	Claim 1; Page 37-39; 50pp; English.
XX	
CC	The invention relates to a novel human protein that shares structural
CC	similarity with animal kinases, including serine-threonine kinases,
CC	particularly Ciron rho-interacting kinases. The proteins of the
CC	invention have neotropic and cytostatic activity. The polynucleotides may
CC	have a use in gene therapy. The encoded novel polypeptides are useful for
CC	generating antibodies, as reagents in diagnostic assays, for identifying
CC	other cellular gene products related to NHP and as reagents in assays for
CC	screening for compounds that are useful in the treatment of mental,
CC	biological or medical disorders and diseases including cancer. The
CC	sequence encodes a novel human kinase of the invention
XX	
SQ	Sequence 6165 BP; 1735 A; 1550 C; 1679 G; 1201 T; 0 U; 0 Other;
	Query Match 86.1%; Score 5661.4; DB 6; Length 6165;
	Best Local Similarity 95.8%; Pred. No. 0;
	Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;
Qy	19 ATGTTGAAGTTCAATATGGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT 78
Db	1 ATGTTGAAGTTCAATATGGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT 60
Qy	79 GCCAGCGCGGCTCCAGGCTCAATCTGTTCTTCCAGGGGAAACACACCCCTTTATGACTCAA 138
Db	61 GCCAGCGGCTCCAGGCTCAATCTGTTCTTCCAGGGGAAACACACCCCTTTATGACTCAA 120
Qy	139 CAGCAGATGCTCTCTTTCCGAGAAGGGATATAGATGGCTCTTTGTCTCTTTGAA 198
Db	121 CAGCAGATGCTCTCTTTCCGAGAAGGGATATAGATGGCTCTTTGTCTCTTTGAA 180
Qy	199 GAATGCACTCAGCTGCTCTCATGAAGATTAAACACGCGAGCAACTTTGTCGGGAAGTAT 258
Db	181 GAA-TGCACTCAGCTGCTCTCATGAAGATTAAACACGCTGAGCAACTTTGTCGGGAAGTAT 240
Qy	259 TCCGACACCATAGCTGAGTTACAGGAGTCTCAGCCTTCGGCAAAGGATTTGGAAGTCTAGA 318
Db	241 TCCGACACCATAGCTGAGTTACAGGAGTCTCAGCCTTCGGCAAAGGATTTGGAAGTCTAGA 300
Qy	319 AGTCTTGTAGTTGGTGCACCTTCTCAGTGCAGGTGGTAAAGAGAGAAACCAACCGG 378
Db	301 AGTCTTGTAGTTGGTGCACCTTCTCAGTGCAGGTGGTAAAGAGAGAAACCAACCGG 360
Qy	379 GACATCTATGCTATGAAAGTGATGAAGAAAGGCTTTATTGGCCCGCAGGAGCAGGTTTCA 438
Db	361 GACATCTATGCTATGAAAGTGATGAAGAAAGGCTTTATTGGCCCGCAGGAGCAGGTTTCA 420
Qy	439 TTTTGTGAGGAAGACCGGAACATATATCTCGAAGCAAGCCCGTGGATCCCCCAATT 498
Db	421 TTTTGTGAGGAAGACCGGAACATATATCTCGAAGCAAGCCCGTGGATCCCCCAATT 480
Qy	499 CAGTATGCTTTTCAGACAATAAATACCTTTTATCTGTAGGAGGAATATCAGGCTCGAGGG 558
Db	481 CAGTATGCTTTTCAGACAATAAATACCTTTTATCTGTAGGAGGAATATCAGGCTCGAGGG 540
Qy	559 GACTTGTGCTACTTTTGAATAGATATAGGACACAGTTAGATGAAAAA CTTGATACAGTTT 618
Db	541 GACTTGTGCTACTTTTGAATAGATATAGGACACAGTTAGATGAAAAA CTTGATACAGTTT 600

Db 3839 -----AGTTTCTCTGCAGTAC 3855
QY 3871 AATGAGCTGAAGCTGGCCCTTGAGAGAGAGAAAGCTCGCTGTGCAGAGCTAGAGAGGCC 3930
Db 3856 AATGAGCTGAAGCTGGCCCTTGAGAGAGAGAAAGCTCGCTGTGCAGAGCTAGAGAGGCC 3915
QY 3931 CTTTCAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCCCAACCGCAAGCA 3990
Db 3916 CTTTCAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCCCAACCGCAAGCA 3975
QY 3991 ACGGACCAACCCACACCCATPCACGGCAGCCACCGCGAGGAGCAGATGCCATGTCCGCC 4050
Db 3976 ACGGACCAACCCACACCCATPCACGGCAGCCACCGCGAGGAGCAGATGCCATGTCCGCC 4035
QY 4051 ATCTGTCGGTCCGAGAGCAGCAGCCAGCTGCATGAGCTCTCGCCCGCCGATCCAGC 4110
Db 4036 ATCTGTCGGTCCGAGAGCAGCAGCCAGCTGCATGAGCTCTCGCCCGCCGATCCAGC 4095
QY 4111 CGCAGAAAGGAGCTTTCAATCCAGAGGAATTTAGTCGGCGCTTTAAGAAAGCAGATGCAC 4170
Db 4096 CGCAGAAAGGAGCTTTCAATCCAGAGGAATTTAGTCGGCGCTTTAAGAAAGCAGATGCAC 4155
QY 4171 CACAAATTTCTCAACCGATTAACGTAGACCTGAACATGCGAGCCACAAAGTGTCTGTG 4230
Db 4156 CACAAATTTCTCAACCGATTTCAACGTAGACCTGAACATGCGAGCCACAAAGTGTCTGTG 4215
QY 4231 TGTCTGGATACCGTGCACCTTGGAGCCAGGCATCCAAATGTCTCGAATGTCTCAGTGTGATG 4290
Db 4216 TGTCTGGATACCGTGCACCTTGGAGCCAGGCATCCAAATGTCTCGAATGTCTCAGTGTGATG 4275
QY 4291 TGTACCCCAAGTGTCTCAAGTGTTCGAGCCAGCTCGGGTTCGCTGTGAATATGCC 4350
Db 4276 TGTACCCCAAGTGTCTCAAGTGTTCGAGCCAGCTCGGGTTCGCTGTGAATATGCC 4335
QY 4351 ACACACTTCCACGAGCCCTTCTGCGTGACAAATGAACTCCCGAGGCTCCAGACCAAG 4410
Db 4336 ACACACTTCCACGAGCCCTTCTGCGTGACAAATGAACTCCCGAGGCTCCAGACCAAG 4395
QY 4411 GAGCCAGCAGCAGCTTGACCTTGAAGGGTGGATGAAGGTCCCGAGGAATTAACAAAGCA 4470
Db 4396 GAGCCAGCAGCAGCTTGACCTTGAAGGGTGGATGAAGGTCCCGAGGAATTAACAAAGCA 4455
QY 4471 GGACACAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4530
Db 4456 GGACACAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4515
QY 4531 GACAAATGAAGCCAGAGAGCTGACAGAGCCGGTGGAAAGATTTGAGCTGCTCTCC 4590
Db 4516 GACAAATGAAGCCAGAGAGCTGACAGAGCCGGTGGAAAGATTTGAGCTGCTCTCC 4575
QY 4591 GACGGGATGTATCTATTGATGTCGGTGGTGTCTTCCGAACTCGCAATACAGCCAA 4650
Db 4576 GACGGGATGTATCTATTGATGTCGGTGGTGTCTTCCGAACTCGCAATACAGCCAA 4635
QY 4651 GCA----- 4653
Db 4636 GCAGATGTCCCATACATCTGAAGATGGAAATCTCACCCGCACACACCTGCTGGCCGGG 4695
QY 4654 ----- 4653
Db 4696 AGAACCCTCTACTTGTAGCTCCAGCTTCCCTGACAAAACAGGGCTGGGTACCCGCTTA 4755
QY 4654 -----GAJAAAGCAGAGCTGATGCTAAACTG 4680
Db 4756 GAATCTAGTTGTCGAGGTGGGAGATTTCTAGGGGAJAAAGCAGAACTGATGCTAAACTG 4815
QY 4681 CTTTGGAACTCCCTGCTGAAACTGGAAGTGTATGACCGTCTAGACATGAACCTGCACGCTG 4740
Db 4816 CTTTGGAACTCCCTGCTGAAACTGGAAGTGTATGACCGTCTAGACATGAACCTGCACGCTG 4875
QY 4741 CCCTTCAGTGAACAGAGTGTGTGTGGGACCCAGAGGAAGGCTCTACGCCCTGGAATGTC 4800
Db 4876 CCCTTCAGTGAACAGAGTGTGTGTGGGACCCAGAGGAAGGCTCTACGCCCTGGAATGTC 4935

QY 4801 TTGAAAACCTCCCTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
Db 4936 TTGAAAACCTCCCTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995
QY 4861 AAGGACCTGAGAGAGCTACTCATGTAGAGAGAGAGCGGCACTGTCTTTGTGGAC 4920
Db 4996 AAGGACCTGAGAGAGCTACTCATGTAGAGAGAGAGCGGCACTGTCTTTGTGGAC 5055
QY 4921 GTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCTCCAGCCGACATCTCA 4980
Db 5056 GTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCTCCAGCCGACATCTCA 5115
QY 4981 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGAGGCAAGATTGAGAAC 5040
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGAGGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTGTCTCATTTCTCGCTACAAACGAAAC 5100
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTGTCTCATTTCTCGCTACAAACGAAAC 5235
QY 5101 CTCAGCAAACTACTGCAATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160
Db 5236 CTCAGCAAACTACTGCAATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5295
QY 5161 TTCAACAAATACAGTATCTCTCATTTGAAACCAATAAATTTCTACGAAATCGACNTGAAGCAG 5220
Db 5296 TTCAACAAATACAGTATCTCTCATTTGAAACCAATAAATTTCTACGAAATCGACNTGAAGCAG 5355
QY 5221 TACAGCTCGAGGAATTTCTTGGATAAGATGACCATTTCTTGGCACCTGTGTCTTTGCC 5280
Db 5356 TACAGCTCGAGGAATTTCTTGGATAAGATGACCATTTCTTGGCACCTGTGTCTTTGCC 5415
QY 5281 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACAGCGAGGGCAGCGAGAG 5340
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACAGCGAGGGCAGCGAGAG 5475
QY 5341 GAGTACTTGTCTGTCTTCCAGATTTGGAGTGTCTGTGATTTCTACGGAAGAGCTAGC 5400
Db 5476 GAGTACTTGTCTGTCTTCCAGATTTGGAGTGTCTGTGATTTCTACGGAAGAGCTAGC 5535
QY 5401 CGCACAGAGATCTCAAGTGGAGTCTTACCTTTTGGCTTTGGCTACAGAGAACCTTAT 5460
Db 5536 CGCACAGAGATCTCAAGTGGAGTCTTACCTTTTGGCTTTGGCTACAGAGAACCTTAT 5595
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTCAGATCCAGGCACTCTCTCAGCA 5520
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTCAGATCCAGGCACTCTCTCAGCA 5655
QY 5521 GGGACCCCTGCGCAGCGTACCTCGAATCCCGAAACCCGGCTACCTGGGCGCTGCCATT 5580
Db 5656 GGGACCCCTGCGCAGCGTACCTCGAATCCCGAAACCCGGCTACCTGGGCGCTGCCATT 5715
QY 5581 TCCTCAGAGAGATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGTGTC 5640
Db 5716 TCCTCAGAGAGATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGTGTC 5775
QY 5641 AAGGAAACCTCTGTAAGGAGTCCGGCACTGAAACACACCGGGGCGCTCCACCTCCCGC 5700
Db 5776 AAGGAAACCTCTGTAAGGAGTCCGGCACTGAAACACACCGGGGCGCTCCACCTCCCGC 5835
QY 5701 AGCAGCCCCAAACAGAGAGGCCCAACCACTGTAACAGAGCAATCAACAGCGGTGCC 5760
Db 5836 AGCAGCCCCAAACAGAGAGGCCCAACCACTGTAACAGAGCAATCAACAGCGGTGCC 5895
QY 5761 TCCAGCCAGCGCCCGAGAGGCCCGAGCCACCGCGAGAGCCCAACACACCCACCGC 5820
Db 5896 TCCAGCCAGCGCCCGAGAGGCCCGAGCCACCGCGAGAGCCCAACACACCCACCGC 5955
QY 5821 TACCGCAGAGGGCGGACCGAGCTGCGAGGCAAGTCTCTCTGGCCGCCCCCTGAGCGA 5880
Db 5956 TACCGCAGAGGGCGGACCGAGCTGCGAGGCAAGTCTCTCTGGCCGCCCCCTGAGCGA 6015

QY 3208 GAAAAAGCGGCAGTGGGAGGCTCGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTT 3267
Db 3241 GAAAAAGCGGCAGTGGGAGGCTCGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTT 3300
QY 3268 GAGTGTCCGGTTCGAGAGCTCGAGAGTCTGGACACCGAGAAACAGACAGCGCGAGA 3327
Db 3301 GAGTGTCCGGTTCGAGAGCTCGAGAGTCTGGACACCGAGAAACAGACAGCGCGAGA 3360
QY 3328 GCCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGTGGAGTGAAGAGGCAAG 3387
Db 3361 GCCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGTGGAGTGAAGAGGCAAG 3420
QY 3388 GCTCAGATTTCTCGTCTCGACGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCCCTC 3447
Db 3421 GCTCAGATTTCTCGTCTCGACGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCCCTC 3480
QY 3448 TCTGACAGGCTCAATGACCTGGAGAGGAGCATGCTATGCTTGAATGAATGCCGAAGC 3507
Db 3481 TCTGACAGGCTCAATGACCTGGAGAGGAGCATGCTATGCTTGAATGAATGCCGAAGC 3540
QY 3508 TTACAGCAGAGCTGGAGCTGACGAGGCTCAACAGAGGCTTCTGGAGAGGCAAGC 3567
Db 3541 TTACAGCAGAGCTGGAGCTGACGAGGCTCAACAGAGGCTTCTGGAGAGGCAAGC 3600
QY 3568 AAATTACAGCAGAGTGGACCTGCGAGAAATCAATTTTCCGTCTGACTCAAGGACTG 3627
Db 3601 AAATTACAGCAGAGTGGACCTGCGAGAAATCAATTTTCCGTCTGACTCAAGGACTG 3660
QY 3628 CAAGAAGCTTAGATCGGCTGATCTACTGAAGCAGAGAGTGAAGTGAAGTATCAG 3687
Db 3661 CAAGAAGCTTAGATCGGCTGATCTACTGAAGCAGAGAGTGAAGTGAAGTATCAG 3720
QY 3688 CTGGAACCAATTCAGGTTCTCTATTTCTCATGAAGAGTGAAGTGAAGGCACTATTTCT 3747
Db 3721 CTGGAACCAATTCAGGTTCTCTATTTCTCATGAAGAGTGAAGTGAAGGCACTATTTCT 3780
QY 3748 CAACAAACCAATTCAGTATGATTTCTGCAAGCAAAATGACCAACCTCTAAAGAAA 3807
Db 3781 CAACAAACCAATTCAGTATGATTTCTGCAAGCAAAATGACCAACCTCTAAAGAAA 3840
QY 3808 AAGGGTTATTTAGTCGACGAGAGGAGCCCTGCTTTTACCCACACAGGTTCTCTCGAG 3867
Db 3841 A-----AGTTTCTCTGCGAG 3855
QY 3868 TACAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGGAA 3927
Db 3856 TACAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGGAA 3915
QY 3928 GCCCTTCAGAGACCCGACATCGAGCTCCGTCGCGCCGGAGGAGAGCTGCCACCGCAA 3987
Db 3916 GCCCTTCAGAGACCCGACATCGAGCTCCGTCGCGCCGGAGGAGAGCTGCCACCGCAA 3975
QY 3988 GCAACGAGACCCACCAACCCATCCAGCCAGCCACCCGCGAGGAGCAGATCGCCATGTCC 4047
Db 3976 GCAACGAGACCCACCAACCCATCCAGCCAGCCACCCGCGAGGAGCAGATCGCCATGTCC 4035
QY 4048 GCCATGTCGGTGGCGCAGAGCAAGCCAGTGCATGAGCCTGCTGGCCCGCCATCC 4107
Db 4036 GCCATGTCGGTGGCGCAGAGCAAGCCAGTGCATGAGCCTGCTGGCCCGCCATCC 4095
QY 4108 AGCGCAGAGAGGCTTCACTCCAGAGGATTTAGTCGCGCTCTTAAGGAAGCGCATG 4167
Db 4096 AGCGCAGAGAGGCTTCACTCCAGAGGATTTAGTCGCGCTCTTAAGGAAGCGCATG 4155
QY 4168 CACCAAAATTTCTCACCGATTCACAGTGAAGTGAACATCGAGGCCACAAAGTGTCT 4227
Db 4156 CACCAAAATTTCTCACCGATTCACAGTGAAGTGAACATCGAGGCCACAAAGTGTCT 4215
QY 4228 GTGTGTCTGGATACCGTGCATTTTGGAGCCAGGCAATCCAAATGTCGAATGTCAGGTG 4287
Db 4216 GTGTGTCTGGATACCGTGCATTTTGGAGCCAGGCAATCCAAATGTCGAATGTCAGGTG 4275
QY 4288 ATGTGTCAACCCCAAGTGTCTCACGCTTTGCCAGCCACCTGCGGCTTGCCTGCTGATAT 4347

Db 4276 ATGTGTCAACCCCAAGTGTCTCACGCTTGCACGCCACCTGCGGCTTGCCTGCTGAAATAT 4335
QY 4348 GCCACACACTTTCACCGAGGCTTTCGCCGTGACAAAATGAATCTCCAGGCTCTCCAGACC 4407
Db 4336 GCCACACACTTTCACCGAGGCTTTCGCCGTGACAAAATGAATCTCCAGGCTCTCCAGACC 4395
QY 4408 AAGGAGCCACAGCAGCAGCTTGCACTCTGGAAGGTGATGAAGGTGCCAGGAATAACAA 4467
Db 4396 AAGGAGCCACAGCAGCAGCTTGCACTCTGGAAGGTGATGAAGGTGCCAGGAATAACAA 4455
QY 4468 CGAGACAGCAAGGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCTCAT 4527
Db 4456 CGAGACAGCAAGGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCTCAT 4515
QY 4528 TATGACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAAGAAATTTGAGCTGTGCTT 4587
Db 4516 TATGACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAAGAAATTTGAGCTGTGCTT 4575
QY 4588 CCCGAGCGGATGATCTATTTCTCATGTCGCTTCCGAACTCGAAATACAGCC 4647
Db 4576 CCCGAGCGGATGATCTATTTCTCATGTCGCTTCCGAACTCGAAATACAGCC 4635
QY 4648 AAAGCAGA----- 4655
Db 4636 AAAGCAGATGTCCCATACATACTGAAGATGGAATCTCACCGCACACCACTGCTGCCCC 4695
QY 4656 ----- 4655
Db 4696 GGGAGAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAAACAGCGCTGGGTACCGCC 4755
QY 4656 -----AAAGCAGAAAGCTGATGCTTAAA 4677
Db 4756 TTAGAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAAGACAGAGCTGATGCTAAA 4815
QY 4678 CTGCTTGGAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACGTGACG 4737
Db 4816 CTGCTTGGAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACGTGACG 4875
QY 4738 CTGCCCCCTCAGTACACAGGTGTTGTTGGTGGCACCGAGGAAGGCTCTACGCCCTGAAT 4797
Db 4876 CTGCCCCCTCAGTACACAGGTGTTGTTGGTGGCACCGAGGAAGGCTCTACGCCCTGAAT 4935
QY 4798 GTCTTGAATACTCCCTAACCCATGTCCAGGAATTTGAGCAGCTTCTTCAAAATTTATAT 4857
Db 4936 GTCTTGAATACTCCCTAACCCATGTCCAGGAATTTGAGCAGCTTCTTCAAAATTTATAT 4995
QY 4858 ATCAAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGCGGCACTGCTCTGTG 4917
Db 4996 ATCAAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGCGGCACTGCTCTGTG 5055
QY 4918 GACGTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATC 4977
Db 5056 GACGTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATC 5115
QY 4978 TCACCCAAATTTTGAAGCTTCAAGGCTGCACTTGTGTTGGGAGGCAAGTGTAG 5037
Db 5116 TCACCCAAATTTTGAAGCTTCAAGGCTGCACTTGTGTTGGGAGGCAAGTGTAG 5175
QY 5038 AACGGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTCTGCTTCTCCGCTACAAACGAA 5097
Db 5176 AACGGGCTCTGCATCTGTGAGCCAATGCCAGCAAAAGTCTGCTTCTCCGCTACAAACGAA 5235
QY 5098 AACCTCAGCAAAATCTGATCTCCGAAAGAGATAGACCTCAGAGCCCTGAGCTGTATC 5157
Db 5236 AACCTCAGCAAAATCTGATCTCCGAAAGAGATAGACCTCAGAGCCCTGAGCTGTATC 5295
QY 5158 CACTTCAACCAATTTACAGTATCTCTTGGAAACCAATAAATTTCTACGAAATCGACATGAAG 5217
Db 5296 CACTTCAACCAATTTACAGTATCTCTTGGAAACCAATAAATTTCTACGAAATCGACATGAAG 5355
QY 5218 CAGTACAGCTCGAGGAATTTCTGGATAAGAAATGAACATTTCTTGTGCACTGCTGTGTTT 5277

Db 5356 CAGTACACGCTCGAGGAATTCCTGGATAGAATGACCATCTCTTGGCACCTGCTGTGTTT 5415
Qy 5278 GCGGCTCTTCAACAGAGCTTCCCTCTCAATCGTCAAGTGAACAGCGAGCGGAGGCA 5337
Db 5416 GCGGCTCTTCAACAGAGCTTCCCTCTCAATCGTCAAGTGAACAGCGAGCGGAGGCA 5475
Qy 5338 GAGGAGTACTGCTGTGTTTCCACGAATTTGAGAGTTCGTGGATTTCTTACGGAAGAGCT 5397
Db 5476 GAGGAGTACTGCTGTGTTTCCACGAATTTGAGAGTTCGTGGATTTCTTACGGAAGAGCT 5535
Qy 5398 AGCCGACAGAGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGGCTACAGAAACCC 5457
Db 5536 AGCCGACAGAGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGGCTACAGAAACCC 5595
Qy 5458 TATCTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACAGCTTCTCA 5517
Db 5596 TATCTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACAGCTTCTCA 5655
Qy 5518 GAGGAGCCCTTGCAGGAGCTACTGTGACATCCCGAACCCCGGTACTGTGGCCCTTGC 5577
Db 5656 GCAGGAGCCCTTGCAGGAGCTACTGTGACATCCCGAACCCCGGTACTGTGGCCCTTGC 5715
Qy 5578 ATTTCCTCAGGAGCGAATTTACTTGGCGTCTCATACAGGATAAATTAAGGTCTCAATTTGC 5637
Db 5716 ATTTCCTCAGGAGCGAATTTACTTGGCGTCTCATACAGGATAAATTAAGGTCTCAATTTGC 5775
Qy 5638 TGCAAGGGAACCTCGTGAAGAGTCCGGCACTGACACACCCGCGGCGCGTCCACCTCC 5697
Db 5776 TGCAAGGGAACCTCGTGAAGAGTCCGGCACTGACACACCCGCGGCGCGTCCACCTCC 5835
Qy 5698 CGCAGCGCCCAACAGCGAGCGCCACCCAGTACACGAGCATCATCACCAGCGCGTG 5757
Db 5836 CGCAGCGCCCAACAGCGAGCGCCACCCAGTACACGAGCATCATCACCAGCGCGTG 5895
Qy 5758 GCTTCAGCCCAAGCGCGCCGGAAGCGCCAGCCAGCGAGAGCGAAGCAGACACCCAC 5817
Db 5896 GCTTCAGCCCAAGCGCGCCGGAAGCGCCAGCCAGCGAGAGCGAAGCAGACACCCAC 5955
Qy 5818 CGTACCGCGAGCGCGGAGCCGCTGTCGCGAGGACAGTCTCTGGCGCCCTCGAG 5877
Db 5956 CGTACCGCGAGCGCGGAGCCGCTGTCGCGAGGACAGTCTCTGGCGCCCTCGAG 6015
Qy 5878 CGAGAGAGTCTCCCGCGCGAGTCTCAGCAGCGGAGAGCGGTCTCCCGGAGGCTG 5937
Db 6016 CGAGAGAGTCTCCCGCGCGAGTCTCAGCAGCGGAGAGCGGTCTCCCGGAGGCTG 6075
Qy 5938 TTGGAAGCAGCAGCAGCGCGCGCTGCTGCGGAGCGCTGAGACCCCGCTGCCAG 5997
Db 6076 TTGGAAGCAGCAGCAGCGCGCGCTGCTGCGGAGCGCTGAGACCCCGCTGCCAG 6135
Qy 5998 GTCAACAGGAGAGGGCA 6017
Db 6136 GTCAACAGGAGAGGGCA 6155

Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease; systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy; stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia; pain; alcoholism; transgenic.

Homo sapiens.

WO200226826-A2.

04-APR-2002.

27-SEP-2001; 2001WO-US042336.

27-SEP-2000; 2000US-0235631P.

27-SEP-2000; 2000US-0235633P.

27-SEP-2000; 2000US-0235808P.

27-SEP-2000; 2000US-0236064P.

27-SEP-2000; 2000US-0236055P.

27-SEP-2000; 2000US-0236066P.

28-SEP-2000; 2000US-0236135P.

03-OCT-2000; 2000US-0237434P.

05-OCT-2000; 2000US-0238321P.

06-OCT-2000; 2000US-0238396P.

06-OCT-2000; 2000US-0238399P.

16-MAR-2001; 2001US-0276667P.

31-MAY-2001; 2001US-0294823P.

12-JUL-2001; 2001US-0304868P.

26-SEP-2001; 2001US-00235631.

(CURA-) CURAGEN CORP.

Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;

Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;

Burgess CB, Padigar M, Kekuda R, Spytek KA, Leach MD, Shimkets RA;

WPI; 2002-499860/53.

P-PSDB; ABG78363.

Novel isolated NOVX polypeptides and polynucleotides homologous to attractin, plexin, papin-like family of proteins, useful for treating atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and stroke.

Claim 8; Page 43-44; 308pp; English.

The invention discloses the isolated human polypeptides, and polynucleotides encoding them, that have been designated NOVX. The polypeptides, polynucleotides and antibodies are useful in treating or preventing a NOVX-associated disorder which is cardiomyopathy, atherosclerosis and diabetes in a human, where the disorder is related to cell signal processing and metabolic pathway modulation. They can also be used in determining the presence of, or predisposition to, a disease associated with altered levels of the polypeptides and polynucleotides of any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for identifying an agent that binds to, or that modulates the expression or activity of the polypeptide, for identifying an agent which is cellular receptor or downstream effector, for treating or preventing a NOVX-associated disorder and as a pharmaceutical composition comprising the polypeptide, polynucleotide or the antibody. The polypeptides and polynucleotides are useful in diagnostic applications (e.g. as a marker for cancerous cells or tissue types) where their amounts are assessed, or for the manufacture of a medicament (e.g. gene therapy) for treating or preventing disorders or syndromes such as hypertension, congenital heart defects, aortic stenosis, obesity, infectious disease, anorexia, cancer, Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders, haemophilia, dyslipidemia, haematopoietic diseases, scleroderma, infertility, idiopathic thrombocytopenic purpura, graft versus host diseases, Crohn's disease, multiple sclerosis, arthritis, psoriasis, disease, systemic lupus erythematosus, asthma, ataxia, autoimmune allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar ataxia, pain and alcoholism. They may also be used as immunogens to produce antibodies specific for the invention, and as vaccines. Transgenic cells containing a NOVX expressing construct are useful to

RESULT 8
ABS63436
ID ABS63436 standard; cDNA; 6189 BP.

XX AC ABS63436;

XX DT 15-NOV-2002 (first entry)

XX DE RHO/RAC-interacting citrin kinase-like human cDNA, designated NOV3b.

XX KW Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; diabetes; cell signalling; metabolic pathway; cellular receptor; downstream effector; cancer; gene therapy; hypertension; congenital heart defect; aortic stenosis; obesity; infectious disease; anorexia; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; haemophilia; dyslipidemia; vaccine; haematopoietic disease; scleroderma; infertility; immunogen; idiopathic thrombocytopenic purpura; graft versus host disease;

CC produce non-human transgenic animals for studying the function and/or
CC activity of the NOVX proteins and for identifying and/or evaluating
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC expressing construct are useful to produce non-human transgenic animals
CC for studying the function and/or activity of the NOVX proteins and for
CC identifying and/or evaluating modulators of NOVX protein activity. The
CC sequences presented in AB563431-AB563444 are the human NOV1-NOV8 cDNAs
XX

Sequence 6139 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;
Query Match 85.7%; Score 5631; DB 6; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;

QY 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 78
DB 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 60

QY 79 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTATGACTCAA 138
DB 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTATGACTCAA 120

QY 139 CAGCAGATGTCCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTTTGAA 198
DB 121 CAGCAGATGTCCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTTTGAA 180

QY 199 GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTCTCCGGAGTAT 258
DB 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTCTCCGGAGTAT 240

QY 259 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGCGAAAGACTTCGAAGTCAGA 318
DB 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGCGAAAGACTTCGAAGTCAGA 300

QY 319 AGTCTGTAGTTGTGGTCACTTTGCTGAAGTCAGTGGTGAAGAGAAAGCAACCGGG 378
DB 301 AGTCTGTAGTTGTGGTCACTTTGCTGAAGTCAGTGGTGAAGAGAAAGCAACCGGG 360

QY 379 GACATCTATCTATGAAAGTGAAGAGAAAGGCTTTATTTGGCCAGGAGCAGGTTTCA 438
DB 361 GACATCTATCTATGAAAGTGAAGAGAAAGGCTTTATTTGGCCAGGAGCAGGTTTCA 420

QY 439 TTTTGTGAGAGAGCGGAAACATATTATCTGAAGCACAAGCCCGTGGATCCCCCAATTA 498
DB 421 TTTTGTGAGAGAGCGGAAACATATTATCTGAAGCACAAGCCCGTGGATCCCCCAATTA 480

QY 499 CAGTATGCCCTTTCCAGGACAAATATCACTTTATCTGATGGAGGAATATCAGCTGGAGG 558
DB 481 CAGTATGCCCTTTCCAGGACAAATATCACTTTATCTGATGGAGGAATATCAGCTGGAGG 540

QY 559 GACTTGCTGCTCACTTTTGAATAGATATGAGCAGTATAGATGAATCACTGATACAGTTT 618
DB 541 GACTTGCTGCTCACTTTTGAATAGATATGAGCAGTATAGATGAATCACTGATACAGTTT 600

QY 619 TACCTAGCTGAGCTGATTTGGCTGTTCCAGCGTTTCATCTGATGGGATACCTGATCGA 678
DB 601 TACCTAGCTGAGCTGATTTGGCTGTTCCAGCGTTTCATCTGATGGGATACCTGATCGG 660

QY 679 GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGATTTT 738
DB 661 GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGATTTT 720

QY 739 GGATCTGCCGGAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTTGGGACC 798
DB 721 GGATCTGCCGGAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTTGGGACC 777

QY 799 CCAGATTACATGCTCTCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGCACTTAC 858
DB 778 CCAGATTACATGCTCTCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGCACTTAC 837

QY 859 GGCCTGGATGTGACTGTGGTGTGAGTGGCGGTGATTTGCCCTATGAGATGATTTATGGGAGA 918
DB 838 GGCCTGGATGTGACTGTGGTGTGAGTGGCGGTGATTTGCCCTATGAGATGATTTATGGGAGA 897

QY 919 TCCCCCTTCGCAGAGGGAAACCTCTGCCAGAACCTTCAATTAACATATGAATTTCCAGCGG 978
DB 898 TCCCCCTTCGCAGAGGGAAACCTCTGCCAGAACCTTCAATTAACATATGAATTTCCAGCGG 957

QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTGAGCAGTGCATTTCTTGATCTGATTCGAAGC 1038
DB 958 TTTTGTGAATTTCCAGATGACCCCAAGTGAGCAGTGCATTTCTTGATCTGATTCGAAGC 1017

QY 1039 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTTGCTGCCATCCTTTCTTC 1098
DB 1018 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTTGCTGCCATCCTTTCTTC 1077

QY 1099 TCTAAATTTGACTCGGAACACATTCGTAACCTCTCTCCCTCCCTCGTTCCACCCCTCAAG 1158
DB 1078 TCTAAATTTGACTCGGAACACATTCGTAACCTCTCTCCCTCCCTCGTTCCACCCCTCAAG 1137

QY 1159 TCTGACGATGACACTCCAAATTTTGTGAACACAGAGAGAAATTCGTGGGTTTCATCCTCT 1218
DB 1138 TCTGACGATGACACTCCAAATTTTGTGAACACAGAGAGAAATTCGTGGGTTTCATCCTCT 1197

QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1278
DB 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1257

QY 1279 TACAGAGGCACTGGGGATTTTGTGTAGATCTGAGTCTGTTGTGCGGCTCTGGACTCC 1338
DB 1258 TACAGAGGCACTGGGGATTTTGTGTAGATCTGAGTCTGTTGTGCGGCTCTGGACTCC 1317

QY 1339 CCTGCCAAGACTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGAGCTTACAAGAC 1398
DB 1318 CCTGCCAAGACTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGAGCTTACAAGAC 1377

QY 1399 TCTCAGAGCAAGTGTCAAGATGGAGCAGAAATGACCCGTTTACATCCGAGAGTGTCA 1458
DB 1378 TCTCAGAGCAAGTGTCAAGATGGAGCAGAAATGACCCGTTTACATCCGAGAGTGTCA 1437

QY 1459 GAGTGGAGGCTGTCTTAGTTCAGAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518
DB 1438 GAGTGGAGGCTGTCTTAGTTCAGAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497

QY 1519 TCCCTCTCGAGCAGGACCTTGCTTACCTACATCACAGAAATGCAAGTCTTAAAGCGAGT 1578
DB 1498 TCCCTCTCGAGCAGGACCTTGCTTACCTACATCACAGAAATGCAAGTCTTAAAGCGAGT 1557

QY 1579 TTGAGCAAGCACGGATGGAGGTGTCCAGAGAGGATGACAAAGCACTGCAAGCTTCTCCAT 1638
DB 1558 TTGAGCAAGCACGGATGGAGGTGTCCAGAGAGGATGACAAAGCACTGCAAGCTTCTCCAT 1617

QY 1639 GATATCAGAGAGCAGAGCCCGAAGCTCCAAAGAAATCAAAGCAGAGGATACCAAGGCTCAA 1698
DB 1618 GATATCAGAGAGCAGAGCCCGAAGCTCCAAAGAAATCAAAGCAGAGGATACCAAGGCTCAA 1677

QY 1699 GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1758
DB 1678 GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1737

QY 1759 CGGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTGCTGAAGAAATCAAG 1818
DB 1738 CGGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTGCTGAAGAAATCAAG 1797

QY 1819 CGGAAGCGCAGAGTCTGAGCATTAATCTTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1878
DB 1798 CGGAAGCGCAGAGTCTGAGCATTAATCTTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1857

QY 1879 GTGGGAGAATATGCAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938
DB 1858 GTGGGAGAATATGCAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1917

QY 1939 CTCCAAGAGAACTGGAGAGGCTG----- 1963
DB 1918 CTCCAAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTGCAG 1977

QY 1964 -----CAGAAGAGCGAGCCGAGAGGGAGCTGGAGAAGCTGGAGAACCAGAGAG 2010
Db 1978 AATATCCGCGACGCAAAAGGAGCGAGCCGAGAGGGAGCTGGAGAAGCTGGAGAACCAGAGAG 2037
QY 2011 GATTCTTCTGAAGGCAATCAGAAAGAGCTGGTGGAGCTGAGAAAGCGCGCATTTCTTG 2070
Db 2038 GATTCTTCTGAAGGCAATCAGAAAGAGCTGGTGGAGCTGAGAAAGCGCGCATTTCTTG 2097
QY 2071 GAGAACAGGTTAAAGAGATAGAGACCATAGAGCCGTAGAGAAACAGACTGAAGATGAC 2130
Db 2098 GAGAACAGGTTAAAGAGATAGAGACCATAGAGCCGTAGAGAAACAGACTGAAGATGAC 2157
QY 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2190
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2217
QY 2191 AAACATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAAGTCCACTGAAACAGAAAGAG 2250
Db 2218 AAACATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAAGTCCACTGAAACAGAAAGAG 2277
QY 2251 CAGCACTATGAGGAAGAAATTAAGTTTGGACAATCAGATAAAGAAACCTGCTGAC 2310
Db 2278 CAGCACTATGAGGAAGAAATTAAGTTTGGACAATCAGATAAAGAAACCTGCTGAC 2337
QY 2311 AAGGAGACACTGGAGAACATGATGCAGACACAGAGGAGGCCCATCAGAAAGGCCAAA 2370
Db 2338 AAGGAGACACTGGAGAACATGATGCAGACACAGAGGAGGCCCATCAGAAAGGCCAAA 2397
QY 2371 ATTCTCAGCGAACAGAAAGCGATGATCAATGCTATGGATTCGAATCCAGATCAGATCCCTGGAA 2430
Db 2398 ATTCTCAGCGAACAGAAAGCGATGATCAATGCTATGGATTCGAATCCAGATCAGATCCCTGGAA 2457
QY 2431 CAGAGATTGTGGAACTGTCTGAAGCCATAAATTCGACGAAATAGCAGTCTTTTACC 2490
Db 2458 CAGAGATTGTGGAACTGTCTGAAGCCATAAATTCGACGAAATAGCAGTCTTTTACC 2517
QY 2491 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550
Db 2518 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGCCCGAGAACCGAAACTGGAGAGCAGCTGGAG 2610
Db 2578 CTGGAGACACAGCTGGGAAGTTGGAGCCCGAGAACCGAAACTGGAGAGCAGCTGGAG 2637
QY 2611 AAGATCAGCCACCAAGACACACAGTGCAGAGAAATCGGCTGCTGAACTGGAGACAAAGATTG 2670
Db 2638 AAGATCAGCCACCAAGACACACAGTGCAGAGAAATCGGCTGCTGAACTGGAGACAAAGATTG 2697
QY 2671 CGGAGGTGAGTCTAGAGACAGAGAGCAGAAACTGGAGCTCAAGGCCAGCTCAGAG 2730
Db 2698 CGGAGGTGAGTCTAGAGACAGAGAGCAGAAACTGGAGCTCAAGGCCAGCTCAGAG 2757
QY 2731 CTACAGCTCTCCTGCGAGGCGGAGTCCAGTTGACAGCCCTGCAGGCTGCACGGCG 2790
Db 2758 CTACAGCTCTCCTGCGAGGCGGAGTCCAGTTGACAGCCCTGCAGGCTGCACGGCG 2817
QY 2791 GGCCTGGAGAGCAGCTTCCGCGGCGAAGACAGAGCTGGAGCGAAAGACACACAGAAAGCT 2850
Db 2818 GGCCTGGAGAGCAGCTTCCGCGGCGAAGACAGAGCTGGAGCGAAAGACACACAGAAAGCT 2877
QY 2851 GAAGAGGATCCAGGCACTCAGGCAATAGATGAAATCCAGGCAAAATTTGATGCT 2910
Db 2878 GAAGAGGATCCAGGCACTCAGGCAATAGATGAAATCCAGGCAAAATTTGATGCT 2937
QY 2911 CTTTCGTAAACAGCTGTACTGTATCAGAGCTGGAGGAGCAGCTAAACAGCTGACCGAG 2970
Db 2938 CTTTCGTAAACAGCTGTACTGTATCAGAGCTGGAGGAGCAGCTAAACAGCTGACCGAG 2997
QY 2971 GACACGCTGAACCTAACAAACAAACTTCTACTTGTCCAAACAACTGATGAGGCTTCT 3030
Db 2998 GACACGCTGAACCTAACAAACAAACTTCTACTTGTCCAAACAACTGATGAGGCTTCT 3057
QY 3031 GGCGCCAAACGACGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTCCGCGCGGAGATC 3090

Db 3058 GGCCCCAAACGACGAGATTGTACAACTCGAAGTGAAGTGAACCACTCTCCGCGGAGATC 3117
QY 3091 ACGGAAACGAGAGATGACAGCTTACCAGCCAGAAACAAACGATGGAGGCTTGAAGACCAACG 3150
Db 3118 ACGGAAACGAGAGATGACAGCTTACCAGCCAGAAACAAACGATGGAGGCTTGAAGACCAACG 3177
QY 3151 TGCACCAATGCTGGAGGAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTAGAA 3210
Db 3178 TGCACCAATGCTGGAGGAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTAGAA 3237
QY 3211 AAAGAGCGGAGTGGAGGGCTGGAGGAGCGTCTCGGTGATGAGAAATCCCAGTTTGGAG 3270
Db 3238 AAAGAGCGGAGTGGAGGGCTGGAGGAGCGTCTCGGTGATGAGAAATCCCAGTTTGGAG 3297
QY 3271 TGTCCGGTTTCGAGAGCTGCGAGAAATGCTGGACACCGAGAAAACAGACGAGGCGAGAGCC 3330
Db 3298 TGTCCGGTTTCGAGAGCTGCGAGAAATGCTGGACACCGAGAAAACAGACGAGGCGAGAGCC 3357
QY 3331 GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGAGTGAAGGAGCACAAGGCT 3390
Db 3358 GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGAGTGAAGGAGCACAAGGCT 3417
QY 3391 GAGATTCTCGTCTGCGAGCGGCTCTCAAAGAGCAGAAAGCTGGAAGGCCGAGAGCTCTCT 3450
Db 3418 GAGATTCTCGTCTGCGAGCGGCTCTCAAAGAGCAGAAAGCTGGAAGGCCGAGAGCTCTCT 3477
QY 3451 GACAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCCGAGCTTA 3510
Db 3478 GACAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCCGAGCTTA 3537
QY 3511 CAGCAGAAGCTGGAGATGGAACGAGACTCAAAACAGAGGCTTCTGGAAGAGCAGACCCAAA 3570
Db 3538 CAGCAGAAGCTGGAGATGGAACGAGACTCAAAACAGAGGCTTCTGGAAGAGCAGACCCAAA 3597
QY 3571 TTAAGCAGCAGATGGAACCTGCGAGAAAATCAATTTTCCGTCTGACTCAAGAGCTGCAA 3630
Db 3598 TTAAGCAGCAGATGGAACCTGCGAGAAAATCAATTTTCCGTCTGACTCAAGAGCTGCAA 3657
QY 3631 GAAGCTTAGATCGGCTGATCTACTCAAGACAGAAAGATGCTTTGGAGTACAGCTG 3690
Db 3658 GAAGCTTAGATCGGCTGATCTACTCAAGACAGAAAGATGCTTTGGAGTACAGCTG 3717
QY 3691 GAAACCAATTCAGGTTCTTATTTCTCATGAAAAGTGAAAATGGAAGGCACATTTTCTCAA 3750
Db 3718 GAAACCAATTCAGGTTCTTATTTCTCATGAAAAGTGAAAATGGAAGGCACATTTTCTCAA 3777
QY 3751 CAAACCAATTCATTTGATTTTCTCAAGCCAAATGGAACCACTGCTTAAAGAGAAAG 3810
Db 3778 CAAACCAATTCATTTGATTTTCTCAAGCCAAATGGAACCACTGCTTAAAGAGAAA - - 3835
QY 3811 GGTTTATTTAGTCGACGGAAGAGGACCTCTTTTACCCACACAGGTTCTCTTGCAGTAC 3870
Db 3836 -----AGGTGCTCTGCGAGTAC 3852
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGTCCGCTGTGCAGAGCTAGAGGAAGCC 3930
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGTCCGCTGTGCAGAGCTAGAGGAAGCC 3912
QY 3931 CTTTCAGAAAGCCCGCATCGAGCTCCGCTCGCGCGGAGGAAGCTGCCACCGCAAGCA 3990
Db 3913 CTTTCAGAAAGCCCGCATCGAGCTCCGCTCGCGCGGAGGAAGCTGCCACCGCAAGCA 3972
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGCAGATGCGCATGCTCGCC 4050
Db 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGCAGATGCGCATGCTCGCC 4032
QY 4051 ATCTGCGGCTGCGCAGAGCAGCCAGTGCATGAGCCTGCTGGCCCGCCCATCCAGC 4110
Db 4033 ATCTGCGGCTGCGCAGAGCAGCCAGTGCATGAGCCTGCTGGCCCGCCCATCCAGC 4092
QY 4111 CGCAGAAAGAGGATCTTCAACTCCAGAGGAATTTAGTCGGGCTCTTTAGGGAACGATGCAC 4170

4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCATGCAC 4152
4171 CACAATATTTCTTCCACCGATTCAACGTAGGACTGAAATCGGAGCCCAAAAGTGTGCTGTG 4230
4153 CACAATATTTCTTCCACCGATTCAACGTAGGACTGAAATCGGAGCCCAAAAGTGTGCTGTG 4212
4231 TGTCTGATACCGTGCACATTTGGACGCCAGGCATCCAAATGTCTCGAATGTCTCAGGTGATG 4290
4213 TGTCTGATACCGTGCACATTTGGACGCCAGGCATCCAAATGTCTAGAAATGTCTCAGGTGATG 4272
4291 TGTACCCCAAGTGTCTCCAGTGTTCAGCAGCACCTCGCGCTTGCCTGCTGAATATGCC 4350
4273 TGTACCCCAAGTGTCTCCAGTGTTCAGCAGCACCTCGCGCTTGCCTGCTGAATATGCC 4332
4351 ACACACTTACCCAGGCGCTTCTCGCGTGAACAAATGAATCTCCACAGGTCTCCAGACCAAG 4410
4333 ACACACTTACCCAGGCGCTTCTCGCGTGAACAAATGAATCTCCACAGGTCTCCAGACCAAG 4392
4411 GAGCCACAGCAGCTTGCACCTTGAAGGTGATGAAGTGCACAGGAATTAACAAACGA 4470
4393 GAGCCACAGCAGCTTGCACCTTGAAGGTGATGAAGTGCACAGGAATTAACAAACGA 4452
4471 GGCACAGAGCTTGGACAGAGTATCATTTCTCGGAGGATCAAAAGTCTCATTTAT 4530
4453 GGCACAGAGCTTGGACAGAGTATCATTTCTCGGAGGATCAAAAGTCTCATTTAT 4512
4531 GACAAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAAAGATTTGAGTGTGCTTCCC 4590
4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAAAGATTTGAGTGTGCTTCCC 4572
4591 GACGGGATGTATCTATTCATGTGTGCGTGTGCTTCCGAACTCCAAATACAGCCAAA 4650
4573 GACGGGATGTATCTATTCATGTGTGCGTGTGCTTCCGAACTCCAAATACAGCCAAA 4632
4651 GCA----- 4653
4633 GCAGATGTCCCATACATACCTGAAGATGGAATCTCACCCGCACACACCTGCTGGCCCGG 4692
4654 ----- 4653
4693 AGAAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCCGCTTA 4752
4654 -----GAAAAAGCAGAGCTGATGCTAAACTG 4680
4753 GAATCAGTTGTCAGAGTGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGCTAAACTG 4812
4681 CTTGGAAACTCTCCCTGTGAAACTTGGAGGTGATGACCGTCTAGACATGAACCTGCAAGTG 4740
4813 CTTGGAAACTCTCCCTGTGAAACTTGGAGGTGATGACCGTCTAGACATGAACCTGCAAGTG 4872
4741 CCTTCAGTGACAGAGTGTGTGGTGGCCACCGAGAGAGGCTCTACCCCTGATGTC 4800
4873 CCTTCAGTGACAGAGTGTGTGGTGGCCACCGAGAGAGGCTCTACCCCTGATGTC 4932
4801 TTGAAAACTCTCCCTAAACCTGTCAGGAAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860
4933 TTGAAAACTCTCCCTAAACCTGTCAGGAAATTTGGAGCAGTCTTCCAAATTTATATTATC 4992
4861 AAGGACTGTGAGAGCTACTCATGTAGCAGGAGAGAGGCGGCTGCTTGTGAGAC 4920
4993 AAGGACTGTGAGAGCTACTCATGTAGCAGGAGAGAGGCGGCTGCTTGTGAGAC 5052
4921 GTGAGAAAGTGAACAGTCTCCCTGGCCAGTCCCACTGCTGCCAGCCGACACTCTCA 4980
5053 GTGAGAAAGTGAACAGTCTCCCTGGCCAGTCCCACTGCTGCCAGCCGACACTCTCA 5112
4981 CCAGATTTTGAAGTGTCAAGGCTGCGCATTTGTTGGGCGAGGCAAGATTGAGAAC 5040
5113 CCAGATTTTGAAGTGTCAAGGCTGCGCATTTGTTGGGCGAGGCAAGATTGAGAAC 5172
5041 GGGCTCTGATCTGTGAGCAGTCCAGCAAGTCTCATTTCTCGCTTACAAACCAAAAC 5100
5173 GGGCTCTGATCTGTGAGCAGTCCAGCAAGTCTCATTTCTCGCTTACAAACCAAAAC 5232

5101 CTCAGCAATATCTCATCCGGAAGAGATAGAGACCTCAGAGCCCTCGAGCTGTATCCAC 5160
5233 CTCAGCAATATCTCATCCGGAAGAGATAGAGACCTCAGAGCCCTCGAGCTGTATCCAC 5292
5161 TTCAACCAATTAACAGTATCTCTCAATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
5293 TTCAACCAATTAACAGTATCTCTCAATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5352
5221 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCAATTTCTTGGACCTGTGTGTTGCC 5280
5353 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCAATTTCTTGGACCTGTGTGTTGCC 5412
5281 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGCAGGAGAG 5340
5413 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGCAGGAGAG 5472
5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5400
5473 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5532
5401 CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCTTTGCTTACAGAGAACCTAT 5460
5533 CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCTTTGCTTACAGAGAACCTAT 5592
5461 CTGTTTGTGACCCACTTCAACTCACTCGAAATTAATGAGATCCAGGACCGCTCTCAGCA 5520
5593 CTGTTTGTGACCCACTTCAACTCACTCGAAATTAATGAGATCCAGGACCGCTCTCAGCA 5652
5591 GGGACCCCTCGCCAGAGTACTTGGACATCCGAAACCGCGCTACCTGGCCCTTGCCTAT 5580
5553 GGGACCCCTCGCCAGAGTACTTGGACATCCGAAACCGCGCTACCTGGCCCTTGCCTAT 5712
5581 TCCTCAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC 5640
5713 TCCTCAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC 5772
5641 AAGGGAACCTCGTGAAGGAGTCGGGCACTGAACACACCGGGCGCTCCACCTCCGC 5700
5773 AAGGGAACCTCGTGAAGGAGTCGGGCACTGAACACACCGGGCGCTCCACCTCCGC 5832
5701 AGCAGCCCAACAGCAGGCGCCACCCACGTAACAAGAGCACATCACCAAGCGCTGGCC 5760
5833 AGCAGCCCAACAGCAGGCGCCACCCACGTAACAAGAGCACATCACCAAGCGCTGGCC 5892
5761 TCCAGCCAGGCGCGCCGAGAGGCGCCAGCCACCCGACACAGAGAGCCAGCACCCCGCCG 5820
5893 TCCAGCCAGGCGCGCCGAGAGGCGCCAGCCACCCGACACAGAGAGCCAGCACCCCGCCG 5952
5821 TACCGCAGGCGCGGAGCCGAGCTGCGAGGACAAAGTCTCTGGCGCCCTCGGAGCGA 5880
5953 TACCGCAGGCGCGGAGCCGAGCTGCGAGGACAAAGTCTCTGGCGCCCTCGGAGCGA 6012
5881 GAGAAGTCTCCCGCCGCTGCTCAGCACGCGAGAGAGCGGTCCCGCGGAGCGTGT 5940
6013 GAGAAGTCTCCCGCCGCTGCTCAGCACGCGAGAGAGCGGTCCCGCGGAGCGTGT 6072
5941 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6000
6073 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6132
6001 AACAGGGAAGAGGCGAGAGTGTCTCAAGTTTTCACGTTTAACTGTCACT 6055
6133 AACAGGTTGGGACCACTTCTCAGTATAAATCTCAGCCAGAAAAAACCAACTCT 6187

RESULT 9
ADA05641
ID ADA05641 standard; cDNA; 6189 BP.
XX
AC ADA05641;
XX
DT 06-NOV-2003 (first entry)


```
XX DE Human NOV1a encoding cDNA SEQ ID NO:1.
XX KW human; NOVX; anti-diabetic; anorectic; antibacterial; virucide;
XX KW immunomodulator; cytostatic; nootropic; neuroprotective;
XX KW antiparkinsonian; antilipemic; gene therapy; human disease;
XX KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX KW immune disorder; haematopoietic disorder; dyslipidemia; gene; ss.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT CDS 1..6162
XX FT /tags a
XX FT /product="NOV1a"
XX PN WO2003029424-A2.
XX PD 10-APR-2003.
XX PP 02-OCT-2002; 2002WO-US031373.
XX PR 02-OCT-2001; 2001US-0326483P.
XX PR 05-OCT-2001; 2001US-0327435P.
XX PR 05-OCT-2001; 2001US-0327449P.
XX PR 09-OCT-2001; 2001US-0327917P.
XX PR 09-OCT-2001; 2001US-0328029P.
XX PR 09-OCT-2001; 2001US-0328044P.
XX PR 09-OCT-2001; 2001US-0328056P.
XX PR 12-OCT-2001; 2001US-0328849P.
XX PR 15-OCT-2001; 2001US-0329414P.
XX PR 17-OCT-2001; 2001US-0330142P.
XX PR 18-OCT-2001; 2001US-0330309P.
XX PR 22-OCT-2001; 2001US-0341058P.
XX PR 24-OCT-2001; 2001US-0339266P.
XX PR 24-OCT-2001; 2001US-0343629P.
XX PR 29-OCT-2001; 2001US-0349575P.
XX PR 01-NOV-2001; 2001US-0346357P.
XX PR 17-APR-2002; 2002US-0373260P.
XX PR 19-APR-2002; 2002US-0373815P.
XX PR 19-APR-2002; 2002US-0373817P.
XX PR 19-APR-2002; 2002US-0373826P.
XX PR 19-APR-2002; 2002US-0373884P.
XX PR 22-APR-2002; 2002US-0374977P.
XX PR 16-MAY-2002; 2002US-0381037P.
XX PR 16-MAY-2002; 2002US-0381038P.
XX PR 16-MAY-2002; 2002US-0381042P.
XX PR 16-MAY-2002; 2002US-0381642P.
XX PR 28-MAY-2002; 2002US-0383656P.
XX PR 29-MAY-2002; 2002US-0383831P.
XX PR 25-JUN-2002; 2002US-0391335P.
XX PR 01-OCT-2002; 2002US-00262511.
XX PA (CURA-) CURAGEN CORP.
XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
XX PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UN;
XX PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
XX PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
XX PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
XX PI Eisen AJ, Gangalli EA, Rieger DK, Spaderna SK;
XX DR WPI; 2003-381626/36.
XX DR P-PSDB; ADA05642.
XX XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
XX PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
XX PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
XX PT pharmacogenomics.
XX XX Claim 20; Page 98-99; 586pp; English.
XX PS
XX PS
```

The present invention describes NOVX proteins, where X can be 1 to 55 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid molecule which encodes a NOVX protein of the invention; (4) a vector comprising the nucleic acid molecule described above; (5) a cell comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity or of latency or predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide. NOVX sequences have anti-diabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian and antilipemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a syndrome associated with a human disease. The polypeptide or the nucleic acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence encodes a human NOVX protein from the present invention.

Sequence 6189 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;

Query Match 85.7%; Score 5631; DB 7; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;

QY 19 ATGTTGAAGTTCAATATGGAGCGCGGAATCTTTGGATGCTGGTGTGCTGAACCCATT 78
DB 1 ATGTTGAAGTTCAATATGGAGCGCGGAATCTTTGGATGCTGGTGTGCTGAACCCATT 60
QY 79 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCTTTATGACTCAA 138
DB 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCTTTATGACTCAA 120
QY 139 CAGCAGATGTCCTCTTCCCGAGAGGGATATTAGATGCCCTCTTGTCTCTTTTCAA 198
DB 121 CAGCAGATGTCCTCTTCCCGAGAGGGATATTAGATGCCCTCTTGTCTCTTTTCAA 180
QY 199 GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCACGTGAGCAACTTTGTCGGAAGTAT 258
DB 181 GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCACGTGAGCAACTTTGTCGGAAGTAT 240
QY 259 TCCGACACCATAGCTAGTACAGGAGCTCCAGCCTTCGGCAAGAGCTTCGAAGTCAGA 318
DB 241 TCCGACACCATAGCTAGTACAGGAGCTCCAGCCTTCGGCAAGAGCTTCGAAGTCAGA 300
QY 319 AGTCTTGTAGGTTCTGTTGCTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAGCAACCGGG 378
DB 301 AGTCTTGTAGGTTCTGTTGCTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAGCAACCGGG 360
QY 379 GACATCTATGCTATGAAGTATGAAGAAGGCTTTATGGCCCGAGGAGCAGTTTCA 438
DB 361 GACATCTATGCTATGAAGTATGAAGAAGGCTTTATGGCCCGAGGAGCAGTTTCA 420
QY 439 TTTTGTGAGGAGCGGACATATTATCTCGAGACAGCCCGTGGATCCCAATTA 498
DB 421 TTTTGTGAGGAGCGGACATATTATCTCGAGACAGCCCGTGGATCCCAATTA 480

Qy	439	CAGTATGCTTTTCAGGACAAAAATCACTTTATCTGTGAGGAGGAATATACGCTTGGAGGG	558
Db	481	CAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGTGTGATGGAATATACGCTTGGAGGG	540
Qy	559	GACTTGCTGTCACTTTTCAATAGATATGAGGACCAAGTTAGATGAAACCTCATACAGTTT	618
Db	541	GACTTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTCATACAGTTT	600
Qy	619	TACCTAGCTGAGCTGAATTTTGGCTGTTTCCACAGCTTCATCTGATGGGATAGCTGCATCGA	678
Db	601	TACCTAGCTGAGCTGAATTTTGGCTGTTTCCACAGCTTCATCTGATGGGATAGCTGCATCGG	660
Qy	679	GACATCAAGCCTGGAACATTTCTGGTTGACCGCACAGACACATCAAGCTGGTGGATTTT	738
Db	661	GACATCAAGCCTGGAACATTTCTGGTTGACCGCACAGACACATCAAGCTGGTGGATTTT	720
Qy	739	GGATCTGCCGCGAAATGAATTTCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC	798
Db	721	GGATCTGCCGCGAAATGAATTTCAAAACA---GGTGAATGCCAAACTCCCGATTGGGACC	777
Qy	799	CCAGATTACATGGCTCTGAAAGTGGCTGACTGTGATGACGGGATGAAAGGACCACTAC	858
Db	778	CCAGATTACATGGCTCTGAAAGTGGCTGACTGTGATGACGGGATGAAAGGACCACTAC	837
Qy	859	GGCTTGGAAGCTGACTGGTGGTCAAGTGGGCGGTGATGGCTATGAGATGATTTATGGGAGA	918
Db	838	GGCTTGGAAGCTGACTGGTGGTCAAGTGGGCGGTGATGGCTATGAGATGATTTATGGGAGA	897
Qy	919	TCCGCCCTTCGAGAGGGAAACCTCTGCCAGAACCTTCAATAACATATGAAATTCACAGGG	978
Db	898	TCCGCCCTTCGAGAGGGAAACCTCTGCCAGAACCTTCAATAACATATGAAATTCACAGGG	957
Qy	979	TTTTTGAATTTTCCAGATGACCCCAAAGTGAGCAGTGACTTTCTTGATCTGATTTCAAGC	1038
Db	958	TTTTTGAATTTTCCAGATGACCCCAAAGTGAGCAGTGACTTTCTTGATCTGATTTCAAGC	1017
Qy	1039	TTGTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC	1098
Db	1018	TTGTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC	1077
Qy	1099	TCTAAATTTGACTGGACACATTCGTAACTCTCTCCTCCCTTCGTTCCCACTCAAG	1158
Db	1078	TCTAAATTTGACTGGACACATTCGTAAAGCTCTCTCCCTTCGTTCCCACTCAAG	1137
Qy	1159	TCGACGATGACACCTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTTCATCTCT	1218
Db	1138	TCGACGATGACACCTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTTCATCTCT	1197
Qy	1219	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG	1278
Db	1198	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG	1257
Qy	1279	TACAGCAAGGCACGTGGGATTTCTTGGTAGATCTCAGTCTGTGTGTGTCGGGTCGTGACTCC	1338
Db	1258	TACAGCAAGGCACGTGGGATTTCTTGGTAGATCTCAGTCTGTGTGTGTCGGGTCGTGACTCC	1317
Qy	1339	CTGCCAAGACTAGCTCCATGGAAAAAATACTTCTCATCAAAACAAGAGCTACAAGAC	1398
Db	1318	CTGCCAAGACTAGCTCCATGGAAAAAATACTTCTCATCAAAACAAGAGCTACAAGAC	1377
Qy	1399	TCTCAGACACAAGTGTCAACAAGTGGAGCAGGAATGACCCGGTTTACATCGAGAGTGTCA	1458
Db	1378	TCTCAGACACAAGTGTCAACAAGTGGAGCAGGAATGACCCGGTTTACATCGAGAGTGTCA	1437
Qy	1459	GAGGTGGAGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAAGA	1518
Db	1438	GAGGTGGAGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAAGA	1497
Qy	1519	TCCTCTCTGGAGCAGGACCTTGCTACCTCATCACAGAAATGCAAGTACTTAAGCCGAGT	1578
Db	1498	TCCTCTCTGGAGCAGGACCTTGCTACCTCATCACAGAAATGCAAGTACTTAAGCCGAGT	1557
Qy	1579	TTGGAGCAAGCACGGATGTGAGGTTGTCGCCAGGAGGATGACAAGCACTGCACTTCTCCAT	1638

Db	1558	TTGAGACAAGC	ACGGAT	TGGAGGTGT	CCACGAGGAGT	GACAAGAC	CTTGCAGCTTCTCCAT	1617
Qy	1639	GATATCAGAGAG	CAGACGCGGAAGCT	CCAGAATAA	CAAGAGCAGGAGT	ACCAGGCTCAA	1698	
Db	1618	GATATCAGAGAG	CAGACGCGGAAGCT	CCAGAATAA	CAAGAGCAGGAGT	ACCAGGCTCAA	1677	
Qy	1699	GTGGAAGAAAT	TGAGTTGAT	GTGAATCA	GTTCGAAGAGGAT	CTTCTCAGCAAGAAGA	1758	
Db	1678	GTGGAAGAAAT	TGAGTTGAT	GTGAATCA	GTTCGAAGAGGAT	CTTCTCAGCAAGAAGA	1737	
Qy	1759	CGGAGTCAT	CTCTACGAAT	CTGAGCTGAGAG	ACTTCGGCTCTC	GCTGCTGAAGATTTCAAG	1818	
Db	1738	CGGAGTCAT	CTCTACGAAT	CTGAGCTGAGAG	ACTTCGGCTCTC	GCTGCTGAAGATTTCAAG	1797	
Qy	1819	CGGAAGCGA	CAGATGT	TCAGCTAA	ACTTGTGAAGGCT	TAAGGATCAAGGAAGCCTGAA	1878	
Db	1798	CGGAAGCGA	CAGATGT	CAGCTAA	ACTTGTGAAGGCT	TAAGGATCAGGGAAGCCTGAA	1857	
Qy	1879	GTGGGAAAT	ATCGGAA	CTGGAGAGAT	CAATGCTGAGCAGCAGCT	CAAAATTCAGGAG	1938	
Db	1858	GTGGGAAAT	ATCGGAA	CTGGAGAGAT	CAATGCTGAGCAGCAGCT	CAAAATTCAGGAG	1917	
Qy	1939	CTCCAAGAGAAA	CTGGAGAGGCTG	-----	-----	-----	1963	
Db	1918	CTCCAAGAGAAA	CTGGAGAGGCTG	TTAAAAG	CCAGCACGAGGCCAC	CGAGCTGTGTCGAG	1977	
Qy	1964	-----	-----	CAAGAG	CCAGCCGAGAGGGAGCT	GGAGNAGCTGCAGAACCCGAGAG	2010	
Db	1978	AAATATCCG	CAGCAAGGAG	CCGAGCGAGGAGCT	GGAGNAGCTGCAGAACCCGAGAG	2037		
Qy	2011	GATTCTTCT	GAAGGCATCAGAAA	GAAGCTGGTGGAA	AGCTTGAGGAAGCCG	CCCATCTCTCTG	2070	
Db	2038	GATTCTTCT	GAAGGCATCAGAAA	GAAGCTGGTGGAA	AGCTTGAGGAAGCCG	CCCATCTCTCTG	2097	
Qy	2071	GAGAACAGT	TAAGAGA	CTAGAGAC	CAATGGAGCT	TAGAGAAAA	CAGACTGAAGGATGAC	2130
Db	2098	GAGAACAGT	TAAGAGA	CTAGAGAC	CAATGGAGCT	TAGAGAAAA	CAGACTGAAGGATGAC	2157
Qy	2131	ATCCAGACAAA	TCCCAA	CAGATCCAGCAGAT	GGCTGTATAAAT	TCTCGAGCTCGAAGAG	2190	
Db	2158	ATCCAGACAAA	TCCCAA	CAGATCCAGCAGAT	GGCTGTATAAAT	TCTCGAGCTCGAAGAG	2217	
Qy	2191	AAACATCGG	AGGCCCAAGTCT	CAGCCCAGCACC	TAGAA	GTGCACTGTAACAGAAAGAG	2250	
Db	2218	AAACATCGG	AGGCCCAAGTCT	CAGCCCAGCACC	TAGAA	GTGCACTGTAACAGAAAGAG	2277	
Qy	2251	CAGCACTAT	GAGAAAAGATTAA	AGTTGGTGGACA	ATCAGATAA	AGNAGACCTGGCTGAC	2310	
Db	2278	CAGCACTAT	GAGAAAAGATTAA	AGTTGGTGGACA	ATCAGATAA	AGNAGACCTGGCTGAC	2337	
Qy	2311	AAGGAGAC	ACTGGAGAACAT	GTGACAGACAC	CGAGGAGGCGCC	CATCAGAAAGGSCAAA	2370	
Db	2338	AAGGAGAC	ACTGGAGAACAT	GTGACAGACAC	CGAGGAGGCGCC	CATCAGAAAGGSCAAA	2397	
Qy	2371	ATTCTCAG	CGAACAGAGGCGAT	GTCAATGTG	ATTGGAAT	TCAGATCAGATCCCTGGAA	2430	
Db	2398	ATTCTCAG	CGAACAGAGGCGAT	GTCAATGTG	ATTGGAAT	TCAGATCAGATCCCTGGAA	2457	
Qy	2431	CAGAGATT	GTGGAACTGTCT	GAAGCCAATA	ACTTTCAGCAGCAAT	TAGCAGTCTTTTACC	2490	
Db	2458	CAGAGATT	GTGGAACTGTCT	GAAGCCAATA	ACTTTCAGCAGCAAT	TAGCAGTCTTTTACC	2517	
Qy	2491	CAAAAGAA	CATGAAGCC	CCAGAGAGAT	GTATTTCTGAA	CTCAGGCCAACAGAAATTTTAC	2550	
Db	2518	CAAAAGAA	CATGAAGCC	CCAGAGAGAT	GTATTTCTGAA	CTCAGGCCAACAGAAATTTTAC	2577	
Qy	2551	CTGGAGACA	CAGCTCGGGAAG	TTGGAGGCC	CCAGAAC	CCGAAAACTGGAGGAGCAGCTGGAG	2610	
Db	2578	CTGGAGACA	CAGCTCGGGAAG	TTGGAGGCC	CCAGAAC	CCGAAAACTGGAGGAGCAGCTGGAG	2637	
Qy	2611	AAGATCAG	CCCAACAGAC	CACAGT	GACAAGAA	TCGGCTGTGGA	ACTGGAGACAAGATTG	2670

Db 2638 AAGATCAGCCACCAAGACCACAGTGCACAAAGAAATCGGCTGCTGGAATCGGAGACAAGATTG 2697
Qy 2671 CGGGAGTTCAGTCTAGACAGCAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2730
Db 2698 CGGGAGTTCAGTCTAGACAGCAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2757
Qy 2731 CTACAGCTCTCCCTGACGAGCGCAGTCCACAGTTGACAGCCCTGACAGCTGCAGGGC 2790
Db 2758 CTACAGCTCTCCCTGACGAGCGCAGTCCACAGTTGACAGCCCTGACAGCTGCAGGGC 2817
Qy 2791 GCCCTGGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACAGCAGAGCT 2850
Db 2818 GCCCTGGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACAGCAGAGCT 2877
Qy 2851 GAAGAGAGATCCAGCACTCAGCGCACATAGATGAAATCCAGCGCAAAATTTGATGCT 2910
Db 2878 GAAGAGAGATCCAGCACTCAGCGCACATAGATGAAATCCAGCGCAAAATTTGATGCT 2937
Qy 2911 CTTCTGAACAGCTGACTGATATCAAGACCTGGAGGACAGCTAAACAGCTGACCGAG 2970
Db 2938 CTTCTGAACAGCTGACTGATATCAAGACCTGGAGGACAGCTAAACAGCTGACCGAG 2997
Qy 2971 GACACGCTGAACTCAACAAACAAACTTCTACTTCTCAAAACAACTCCGATCAGGCTTCT 3030
Db 2998 GACACGCTGAACTCAACAAACAAACTTCTACTTCTCAAAACAACTCCGATCAGGCTTCT 3057
Qy 3031 GGCGCAACGACGAGATTGTAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3090
Db 3058 GGCGCAACGACGAGATTGTAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3117
Qy 3091 ACGGACGAGATGAGCTTACCAGCCAGAGCAACAGATGGAGCTCTGAAGACCCAG 3150
Db 3118 ACGGACGAGATGAGCTTACCAGCCAGAGCAACAGATGGAGCTCTGAAGACCCAG 3177
Qy 3151 TGCACCATGCTGGAGAAACAGTCTATGATTTGGAGCCCTAAACGATGAGCTGTAGAA 3210
Db 3178 TGCACCATGCTGGAGAAACAGTCTATGATTTGGAGCCCTAAACGATGAGCTGTAGAA 3237
Qy 3211 AAAGACGGAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTGAG 3270
Db 3238 AAAGACGGAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTGAG 3297
Qy 3271 TGTTCGGGTTTCAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGAGGCGAGAGCC 3330
Db 3298 TGTTCGGGTTTCAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGAGGCGAGAGCC 3357
Qy 3331 GATCAGCGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGCACAGGCT 3390
Db 3358 GATCAGCGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGCACAGGCT 3417
Qy 3391 GAGATTCTCGCTCGCAGCAGGCTCTCAAGAGCAGAGCTGAAGCGCAGAGCCCTCTCT 3450
Db 3418 GAGATTCTCGCTCGCAGCAGGCTCTCAAGAGCAGAGCTGAAGCGCAGAGCCCTCTCT 3477
Qy 3451 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCGAAAGCTTA 3510
Db 3478 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCGAAAGCTTA 3537
Qy 3511 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCGGAAGAGCAAGCCAAA 3570
Db 3538 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCGGAAGAGCAAGCCAAA 3597
Qy 3571 TTACAGCAGCAGATGGACCTCGCAGAAAATCAGATTTCCGCTGACTCAAGGACTGCAA 3630
Db 3598 TTACAGCAGCAGATGGACCTCGCAGAAAATCAGATTTCCGCTGACTCAAGGACTGCAA 3657
Qy 3631 GAAGCTCTAGATCGGCTGATCTACTGGAAGACAGAAAGAGTGACTTTGAGTATCAGCTG 3690
Db 3658 GAAGCTCTAGATCGGCTGATCTACTGGAAGACAGAAAGAGTGACTTTGAGTATCAGCTG 3717
Qy 3691 GAAAACATTCAGCTTCTATTTCTCATGAAAAGTGAAGGCAAGCTTATTTCTCAA 3750
Db 3718 GAAAACATTCAGCTTCTATTTCTCATGAAAAGTGAAGGCAAGCTTATTTCTCAA 3777

Qy 3751 CAAACCAAACTCATTGATTTTCTGAAAGCCAAATGACCAACCTGCTAAAAAGAAAAG 3810
Db 3778 CAAACCAAACTCATTGATTTTCTGAAAGCCAAATGACCAACCTGCTAAAAAGAAA-- 3835
Qy 3811 GGTTTATTTAGTCAGCGAAAGAGACCTTGCTTTACCCACACAGGTTCTCTCAGTAC 3870
Db 3836 -----AGGTGCTCTCAGTAC 3852
Qy 3871 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCC 3930
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCC 3912
Qy 3931 CTTCAGAAAGCCCATCGAGCTCCGCTCCGCCCGGAGGAAGTGCACACCGAAAGCA 3990
Db 3913 CTTCAGAAAGCCCATCGAGCTCCGCTCCGCCCGGAGGAAGTGCACACCGAAAGCA 3972
Qy 3991 ACGGACCAACACACCCATCCAGCCAGCCAGTGCATGAGCCCTGCTGGCCCGCCATCCAGC 4050
Db 3973 ACGGACCAACACACCCATCCAGCCAGCCAGTGCATGAGCCCTGCTGGCCCGCCATCCAGC 4032
Qy 4051 ATCGTGGGCTGCCAGAGCACCCAGTGCATGAGCCCTGCTGGCCCGCCATCCAGC 4110
Db 4033 ATCGTGGGCTGCCAGAGCACCCAGTGCATGAGCCCTGCTGGCCCGCCATCCAGC 4092
Qy 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAGAAACGATGCAC 4170
Db 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAGAAACGATGCAC 4152
Qy 4171 CACAATATTTCTCAACGATTCAGCTAGGACTGAAATGCGAGGCAACAAGTGTGCTGTG 4230
Db 4153 CACAATATTTCTCAACGATTCAGCTAGGACTGAAATGCGAGGCAACAAGTGTGCTGTG 4212
Qy 4231 TGTCTGATACCGTGCACTTTGGAGCGCCAGGATCCAAATGCTCAGAAATGTCAGGTGATG 4290
Db 4213 TGTCTGATACCGTGCACTTTGGAGCGCCAGGATCCAAATGCTCAGAAATGTCAGGTGATG 4272
Qy 4291 TGTCAACCCAGTGTCCAGTGTCTGCCAGCACCTCGGCTTGCCTGCTGAATATGTC 4350
Db 4273 TGTCAACCCAGTGTCCAGTGTCTGCCAGCACCTCGGCTTGCCTGCTGAATATGTC 4332
Qy 4351 ACACATTTCAACCGAGGCTTTCGCGTGACAAAATGAATGAACTCCCAAGTCTCCAGACCAAG 4410
Db 4333 ACACATTTCAACCGAGGCTTTCGCGTGACAAAATGAATGAACTCCCAAGTCTCCAGACCAAG 4392
Qy 4411 GAGCCAGCAGCAGCTTGACCTTGAAGGTTGATGAGTGGCCCGCCAGGAATTAACAACGA 4470
Db 4393 GAGCCAGCAGCAGCTTGACCTTGAAGGTTGATGAGTGGCCCGCCAGGAATTAACAACGA 4452
Qy 4471 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTCGAGGGATCAAAAGTCTCTCATTTAT 4530
Db 4453 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTCGAGGGATCAAAAGTCTCTCATTTAT 4512
Qy 4531 GACAAATGAAGCCAGAGAGCTGGAAGGCGGTTGGAAGAAATTTGAGCTGTGCTTCCC 4590
Db 4513 GACAAATGAAGCCAGAGAGCTGGAAGGCGGTTGGAAGAAATTTGAGCTGTGCTTCCC 4572
Qy 4591 GACGGGATGATCTATTCTATGTTGGCTTCCGAACTCGCAATACAGCCAAA 4650
Db 4573 GACGGGATGATCTATTCTATGTTGGCTTCCGAACTCGCAATACAGCCAAA 4632
Qy 4651 GCA----- 4653
Db 4633 GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACCACTGCTGGCCCGGG 4692
Qy 4654 ----- 4653
Db 4693 AGAAACCTCTACTTGTAGTCCAGCTTCCCTGACAAAACAGCGCTGGGTACCGCCCTTA 4752
Qy 4654 -----GAAAAGCAGAAAGCTGATGCTAAACTG 4680
Db 4753 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAANAGCAGAGCTGATGCTAACTG 4812

Qy	4681	CTTGGAAACTCCCTGCTGTAAGAACTCGGAAGGTGATGACCGTCTAGACATGAACTGCACGCTG	474
Db	4813	CTTGGAAACTCCCTGCTGTAAGAACTCGGAAGGTGATGACCGTCTAGACATGAACTGCACGCTG	4872
Qy	4741	CCCTTCAGTGACCAAGGTGGTGGTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGAAATGTC	4800
Db	4873	CCCTTCAGTGACCAAGGTGGTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGAAATGTC	4832
Qy	4801	TTGAAAAACTCCCTAAACCCCATGTCCCAGGAATTTGGACAGTCTTCTCCAAATTTATATTATC	4860
Db	4933	TTGAAAAACTCCCTAAACCCCATGTCCCAGGAATTTGGACAGTCTTCTCCAAATTTATATTATC	4992
Qy	4861	ANGAACCTGGAGAGCTACTCATGATAGCAGGAGAAAGCGGCACATGTGTCTTTGTGGAC	4920
Db	4993	AAGAACCTGGAGAGCTACTCATGATAGCAGGTAAGAGCGGCACATGTGTCTTTGTGGAC	5052
Qy	4921	GTGAAGAAAGTGAACCAAGTCCTCTGGCCAGTCGCCACCTGCTGCTGCCAGCCGACATCTCA	4980
Db	5053	GTGAAGAAAGTGAACCAAGTCCTCTGGCCAGTCGCCACCTGCTGCTGCCAGCCGACATCTCA	5112
Qy	4981	CCCAACAATTTTGAAGCTGTCAAGGGCTGCCAATTGTGTTTGGGGAGCAAGATGTGAAC	5040
Db	5113	CCCAACAATTTTGAAGCTGTCAAGGGCTGCCAATTGTGTTTGGGGAGCAAGATGTGAAC	5172
Qy	5041	GGGCTCTGCAATCTGTGCAGCATGCCAGCAAAAGTCGTCAATCTCOCGTACAAACGAAAC	5100
Db	5173	GGGCTCTGCAATCTGTGCAGCATGCCAGCAAAAGTCGTCAATCTCOCGTACAAACGAAAC	5232
Qy	5101	CTCAGCAAAATATGTCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5160
Db	5233	CTCAGCAAAATATGTCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5292
Qy	5161	TTCCACCAATTTACAGTATCCCTCATTTGGAACCAATAATTTCTACCAAAATCGACATGAGCAG	5220
Db	5293	TTCCACCAATTTACAGTATCCCTCATTTGGAACCAATAATTTCTACCAAAATCGACATGAGCAG	5352
Qy	5221	TACACGCTCGAGGAATTCCTGGATAAGAAATGACCATTCCTTTGGCACCGTGTGTTTGGCC	5280
Db	5353	TACACGCTCGAGGAATTCCTGGATAAGAAATGACCATTCCTTTGGCACCGTGTGTTTGGCC	5412
Qy	5281	GCCTCTTCCACAGCTTCCTGTCTCAATCGTGCAGGTGAACGCGCAGGGCAGCAGAG	5340
Db	5413	GCCTCTTCCACAGCTTCCTGTCTCAATCGTGCAGGTGAACGCGCAGGGCAGCAGAG	5472
Qy	5341	GAGTACTGTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTACGGAACACGTAGC	5400
Db	5473	GAGTACTGTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTACGGAACACGTAGC	5532
Qy	5401	CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTTGGCCCTTTGCCTACAGAGAACCCCTAT	5460
Db	5533	CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTTGGCCCTTTGCCTACAGAGAACCCCTAT	5592
Qy	5461	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGCAGCAGCTCCTCAGCA	5520
Db	5593	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGCAGCAGCTCCTCAGCA	5652
Qy	5521	GGGACCCCTGCCGAGCGTACCTTGGACATCCCGAAACCCGCGCTACCTGGGCCCTGCCATT	5580
Db	5653	GGGACCCCTGCCGAGCGTACCTTGGACATCCCGAAACCCGCGCTACCTGGGCCCTGCCATT	5712
Qy	5581	TTCTCAGGAGCGATTTACTTTGGCGTCTCATACGAGTAAATTAAGGTCATTTGCTCG	5640
Db	5713	TTCTCAGGAGCGATTTACTTTGGCGTCTCATACGAGTAAATTAAGGTCATTTGCTCG	5772
Qy	5641	AAGGGAAACCTCGTGAAGAGTCCGCACTGAAACCAACCGGGGCCCGTCCACCTCCCGC	5700
Db	5773	AAGGGAAACCTCGTGAAGAGTCCGCACTGAAACCAACCGGGGCCCGTCCACCTCCCGC	5832
Qy	5701	AGCAGCCCAACAGCGAGGCCACCAAGTACAAAGACGACATCAACGAGCGGTGGCC	5760
Db	5833	AGCAGCCCAACAGCGAGGCCACCAAGTACAAAGACGACATCAACGAGCGGTGGCC	5892
Qy	5761	TCCAGCCCAAGCGCGGCCGAAGCGCCAGCCACCCGCGAGAGCAACCCACCCG	5820

Db	5893	TCCAGCCGAGCGCCGCCGAAGGCCCCAGGCACCCGCGAGAGCGCAACACACCCACCGC	5952
Qy	5821	TACCGCGAGGGCGGACCCGAGCTCGCGAGGACAAGTCTCTGTGCGCCGCCCTTGGAGCGA	5880
Db	5953	TACCGCGAGGGCGGACCGAGCTCGCGAGGACAAGTCTCTGTGCGCCGCCCTTGGAGCGA	6012
Qy	5881	GAGAAGTCCCGCGCCGATGCTCAGCACCGCGAGAGCGGTCCCGCGGAGCGCTGTTT	5940
Db	6013	GAGAAGTCCCGCGCCGATGCTCAGCACCGCGAGAGCGGTCCCGCGGAGCGCTGTTT	6072
Qy	5941	GAGACAGCAGCAGGGCCCGCTCCCTGCGGGAGCCGTGAGGACCCCGCTGCCAGTG	6000
Db	6073	GAGACAGCAGCAGGGCCCGCTCCCTGCGGGAGCCGTGAGGACCCCGCTGCCAGTG	6132
Qy	6001	AACAAGGGAAGAGGGCAGAGTGCCTCTCAAGTTTTTTCACGGTTAACACTGTCACT	6055
Db	6133	AACAAGGTGGGACCACTTCTCAGTATATAATCTCAGCCAGAAAAACCAACTCCT	6187
RESULT 10			
ABS63435			
XX	ID	ABS63435 standard; cDNA; 6201 BP.	
XX	AC	ABS63435;	
XX	XX		
DT	15-NOV-2002	(first entry)	
XX	XX	Human cDNA, homologous to kinases, designated NOV3a.	
DE	XX		
XX	XX	Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;	
KW	XX	atherosclerosis; diabetes; cell signalling; metabolic pathway;	
KW	XX	cellular receptor; downstream effector; cancer; gene therapy;	
KW	XX	hypertension; congenital heart defect; aortic stenosis; obesity;	
KW	XX	infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;	
KW	XX	neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;	
KW	XX	haematopoietic disease; scleroderma; fertility; immunogen;	
KW	XX	idiopathic thrombocytopenic purpura; graft versus host disease;	
KW	XX	Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;	
KW	XX	systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;	
KW	XX	stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;	
XX	XX	pain; alcoholism; transgenic.	
XX	XX		
OS	XX	homo sapiens.	
XX	XX		
FN	XX	WO200226826-A2.	
XX	XX		
PD	04-APR-2002.		
XX	XX		
PF	27-SEP-2001;	2001WO-US042336.	
XX	XX		
PR	27-SEP-2000;	2000US-0235631P.	
XX	XX		
PR	27-SEP-2000;	2000US-0235633P.	
XX	XX		
PR	27-SEP-2000;	2000US-0235808P.	
XX	XX		
PR	27-SEP-2000;	2000US-0236064P.	
XX	XX		
PR	27-SEP-2000;	2000US-0236065P.	
XX	XX		
PR	28-SEP-2000;	2000US-0236135P.	
XX	XX		
PR	03-OCT-2000;	2000US-0237434P.	
XX	XX		
PR	05-OCT-2000;	2000US-0238321P.	
XX	XX		
PR	06-OCT-2000;	2000US-0238396P.	
XX	XX		
PR	08-OCT-2000;	2000US-0238399P.	
XX	XX		
PR	16-MAR-2001;	2001US-0276667P.	
XX	XX		
PR	31-MAY-2001;	2001US-0294823P.	
XX	XX		
PR	12-JUL-2001;	2001US-0304868P.	
XX	XX		
PR	26-SEP-2001;	2001US-00235631.	
XX	XX		
PA	(CURA-) CURAGEN CORP.		
XX	XX		
PI	Gerlach VL,	MacDougall JR, Smithson G, Millet I, Stone D;	
XX	Guinther E,	Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;	
PI	Burgess CE,	Padigar M, Kekuda R, Spytak KA, Leach MD, Shimkets RA;	
XX	XX		

DR WPI; 2002-499860/53.
DR P-PSDB; ABG78362.
XX Novel isolated NOVX polypeptides and polynucleotides homologous to
PT attractin, plexin, papin-like family of proteins, useful for treating
PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and
PT stroke.
XX
XX
XX Claim 8; Page 40-42; 308pp; English.
CC The invention discloses the isolated human polypeptides, and
CC polynucleotides encoding them, that have been designated NOVX. The
CC polypeptides, polynucleotides and antibodies are useful in treating or
CC preventing a NOVX-associated disorder which is cardiomyopathy,
CC atherosclerosis and diabetes in a human, where the disorder is related to
CC cell signal processing and metabolic pathway modulation. They can also be
CC used in determining the presence of, or predisposition to, a disease
CC associated with altered levels of the polypeptides and polynucleotides of
CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for
CC identifying an agent that binds to, or that modulates the expression or
CC activity of the polypeptide, for identifying an agent which is cellular
CC receptor or downstream effector, for treating or preventing a NOVX-
CC associated disorder and as a pharmaceutical composition comprising the
CC polypeptide, polynucleotide or the antibody. The polypeptides and
CC polynucleotides are useful in diagnostic applications (e.g. as a marker
CC for cancerous cells or tissue types) where their amounts are assessed, or
CC for the manufacture of a medicament (e.g. gene therapy) for treating or
CC preventing disorders or syndromes such as hypertension, congenital heart
CC defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,
CC Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,
CC haemophilia, dyslipidemia, haematopoietic diseases, scleroderma,
CC fertility, idiopathic thrombocytopenic purpura, graft versus host
CC diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
CC allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar
CC ataxia, pain and alcoholism. They may also be used as immunogens to
CC produce antibodies specific for the invention, and as vaccines.
CC Transgenic cells containing a NOVX expressing construct are useful to
CC produce non-human transgenic animals for studying the function and/or
CC activity of the NOVX proteins and for identifying and/or evaluating
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC expressing construct are useful to produce non-human transgenic animals
CC for studying the function and/or activity of the NOVX proteins and for
CC identifying and/or evaluating modulators of NOVX protein activity. The
CC sequences presented in ABS63431-ABS63444 are the human NOV1-NOV8 cDNAs
XX
SQ Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;
Query Match 85.6%; Score 5629.8; DB 5; Length 6201;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;
19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGTAACCCATT 78
1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGTAACCCATT 60
79 GCCAGCGCGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCCCTTTATGACTCAA 138
61 GCCAGCGCGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCCCTTTATGACTCAA 120
139 CAGCAGATGTCCTCTTTCCCGAGAGAGGATATTAGATGCCCTTTGTTCTTTTGAA 198
121 CAGCAGATGTCCTCTTTCCCGAGAGAGGATATTAGATGCCCTTTGTTCTTTTGAA 180
199 GAATGCAGTCAGCTGCTCTGATGAGATTAAGCAGCTGAGCAACTTTGTCGGAGATAT 258
181 GAATGCAGTCAGCTGCTCTGATGAGATTAAGCAGCTGAGCAACTTTGTCGGAGATGT 240
259 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCAAAGGACTTCGAGTCA 318
241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCAAAGGACTTCGAGTCA 300
319 AGTCTTGAGTGTGGTCTACTTTGCTGAAGTCGAGTGGTGAAGAGAAAGCAACCGGG 378

301 AGTCTTGAGTGTGGTCTACTTTTGTCTCAAGTGCAGTGGTAAAGAGAGAAAGCAACCGGG 360
379 GACATCTATGCTATGAAAGTGAAGAGAGAGCTTTATTTGGCCAGGAGCAGGTTTCA 438
361 GACATCTATGCTATGAAAGTGAAGAGAGAGCTTTATTTGGCCAGGAGCAGGTTTCA 420
439 TTTTGGAGAGAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 498
421 TTTTGGAGAGAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480
499 CAGTATGCTTTCAGGACAAATAACCTTATCTGATGGAGGAATATCAGCCTGGAGGG 558
481 CAGTATGCTTTCAGGACAAATAACCTTATCTGATGGAGGAATATCAGCCTGGAGGG 540
559 GACTTGTCTGCTCACTTTTGAATAGATATGAGGACAGTATAGATGAAACCTGATACAGTTT 618
541 GACTTGTCTGCTCACTTTTGAATAGATATGAGGACAGTATAGATGAAACCTGATACAGTTT 600
619 TACCTAGCTAGCTGATTTTGGCTGTTCAAGCGTTCATCTGATGGGATAGCTGATCGA 678
601 TACCTAGCTAGCTGATTTTGGCTGTTCAAGCGTTCATCTGATGGGATAGCTGATCGA 660
679 GACATCAAGCCTGAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 738
661 GACATCAAGCCTGAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 720
739 GGATCTGCCGCGAAATGAATTCAAAACAAGATGTGTAATGCCAACTCCCGATTTGGGACC 798
721 GGATCTGCCGCGAAATGAATTCAAAACAAGATGTGTAATGCCAACTCCCGATTTGGGACC 777
799 CCAGATTAATGCTCTGCTGAGTCTGCTGATGAAAGCGGGATGAAAGAGGACCTTAC 858
778 CCAGATTAATGCTCTGCTGAGTCTGCTGATGAAAGCGGGATGAAAGAGGACCTTAC 837
859 GGCCTGGACTGTGACTGTGTGTGCTGAGTGGCGGTGATTCCTATGAGATGATTTATGGGAGA 918
838 GGCCTGGACTGTGACTGTGTGTGCTGAGTGGCGGTGATTCCTATGAGATGATTTATGGGAGA 897
919 TCCGCTTCGAGAGGAGAACCTCTCCAGAACCTTCAATAACATTAATTAATTTCCAGCGG 978
898 TCCGCTTCGAGAGGAGAACCTCTCCAGAACCTTCAATAACATTAATTAATTTCCAGCGG 957
979 TTTTGGAAATTTCCAGATGACCCCAAGATGAGCAGTACTTCTTGATCTGATTTCAAAGC 1038
958 TTTTGGAAATTTCCAGATGACCCCAAGATGAGCAGTACTTCTTGATCTGATTTCAAAGC 1017
1039 TTGTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTTTGCTGCGCATCCTTTCTTC 1098
1018 TTGTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTTTGCTGCGCATCCTTTCTTC 1077
1099 TCTAAATTTGACTGGAACAAACATTCGTAACTCTCTCCGCTTCTGTTCCACCCCTCAAG 1158
1078 TCTAAATTTGACTGGAACAAACATTCGTAACTCTCTCCGCTTCTGTTCCACCCCTCAAG 1137
1159 TCTGACGATGACACTCCAAATTTTGAATGAAACAGAGAGAAATTCGTGGGTTTCACTCT 1218
1138 TCTGACGATGACACTCCAAATTTTGAATGAAACAGAGAGAAATTCGTGGGTTTCACTCT 1197
1219 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1278
1198 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1257
1279 TACAGCAAGGCACTGGGATTTCTTGGTGAATCTGAGTCTGTTGTGTCGGGTCGAGCTCC 1338
1258 TACAGCAAGGCACTGGGATTTCTTGGTGAATCTGAGTCTGTTGTGTCGGGTCGAGCTCC 1317
1339 CCGTGAAGCTAGCTCCATCGGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1398
1318 CCGTGAAGCTAGCTCCATCGGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1377
1399 TCTCAGGCAAGTGTCACAAGATGGAGGAAATGACCCGTTTACATCGGAGAGTGTCA 1458

Db 1378 TCTCAGGACAAAGTGCTCA CAAGATGGAGCGAGAAATGACCCGGTTACATCGGAGAGTGCTCA 1437
Qy 1459 GAGGTGGAGGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCTCTGAGAGCTCAGAGA 1518
Db 1438 GAGGTGGAGGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCTCTGAGAGCTCAGAGA 1497
Qy 1519 TCCCTCTCTGGAGCAGGACCTTGCTTACCTACATCA CAGAAATGCAAGTAAAGGCAAGT 1578
Db 1498 TCCCTCTCTGGAGCAGGACCTTGCTTACCTACATCA CAGAAATGCAAGTAAAGGCAAGT 1557
Qy 1579 TTGAGCAAGCAGGATGGAGGTGTCCAGGAGGTGACAAAGCACTGCAAGCTTCTCCAT 1638
Db 1558 TTGAGCAAGCAGGATGGAGGTGTCCAGGAGGTGACAAAGCACTGCAAGCTTCTCCAT 1617
Qy 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAAGGCTCAA 1698
Db 1618 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAAGGCTCAA 1677
Qy 1699 GTGCAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1758
Db 1678 GTGCAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1737
Qy 1759 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATCAAAG 1818
Db 1738 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATCAAAG 1797
Qy 1819 CGGAAAGCGCAGAGATGTCAAGTAAATCTTTGAAAGCTTAAGGATCAAGGAAAGCTGAA 1878
Db 1798 CGGAAAGCGCAGAGATGTCAAGTAAATCTTTGAAAGCTTAAGGATCAAGGAAAGCTGAA 1857
Qy 1879 GTGGCAATATGCGAACTTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTTCAAGGAG 1938
Db 1858 GTGGCAATATGCGAACTTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTTCAAGGAG 1917
Qy 1939 CTCAGAGAAAATGGAGAAAGGCT----- 1962
Db 1918 CTCAGAGAAAATGGAGAAAGGCTGTAAGAGCCAGCAGCAGGAGCCACCGAGCTGTGCGAG 1977
Qy 1963 -----GCAGAGAGCGAGCGGAGAGGCTGGAGAGCTGAGAGCTGAGAGCGAGAG 2010
Db 1978 AATATCCGCGCAGGCAAGAGGCGAGCGGAGGAGCTGGAGAGCTGAGAGCGAGAG 2037
Qy 2011 GATTTCTTGAAGGCTCAGAAAAGAGCTGGTGGAAAGCTGAGGAAACCGCCGCTTCTCTG 2070
Db 2038 GATTTCTTGAAGGCTCAGAAAAGAGCTGGTGGAAAGCTGAGGAAACCGCCGCTTCTCTG 2097
Qy 2071 GAGAAACAGGTAAGAGACTAGAGACATGAGAGCCATGAGAGCTAGAGAAAACAGATGAAGTAGAC 2130
Db 2098 GAGAAACAGGTAAGAGACTAGAGACATGAGAGCCATGAGAGCTAGAGAAAACAGATGAAGTAGAC 2157
Qy 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2190
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2217
Qy 2191 AATCATCGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAAAGTGCACCTGAAAACAGAAAAGAG 2250
Db 2218 AATCATCGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAAAGTGCACCTGAAAACAGAAAAGAG 2277
Qy 2251 CAGCAGCTATGAGGAAAGATTTAAGTGTGGCAATCAGATTAAGAAAGACCTGGCTGAC 2310
Db 2278 CAGCAGCTATGAGGAAAGATTTAAGTGTGGCAATCAGATTAAGAAAGACCTGGCTGAC 2337
Qy 2311 AAGGAGACACTGAGAAACATGATGAGAGACAGAGGAGGAGCCCATGAGAGGAGCAAA 2370
Db 2338 AAGGAGACACTGAGAAACATGATGAGAGACAGAGGAGGAGGAGCCCATGAGAGGAGCAAA 2397
Qy 2371 ATTCTCAGGAAACAGAGGAGTATCAATGCTATGCTATGCTTCAAGATCAGATCCCTGGAA 2430
Db 2398 ATTCTCAGGAAACAGAGGAGTATCAATGCTATGCTATGCTTCAAGATCAGATCCCTGGAA 2457
Qy 2431 CAGAGGATTTGGAAGCTGTGAGGCAATAACTTTGAGCAAAATAGCAGTCTTTTACC 2490
Db 2458 CAGAGGATTTGGAAGCTGTGAGGCAATAACTTTGAGCAAAATAGCAGTCTTTTACC 2517

Qy 2491 CAAAGGAAACATGAAGCGCCAAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAAATTTTAC 2550
Db 2518 CAAAGGAAACATGAAGCGCCAAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAATTTTAC 2577
Qy 2551 CTGGAGACACAGGCTGGGAAAGTTGGAGGCCCAGAAACCGAAAACCTGGAGGACAGCTGGAG 2610
Db 2578 CTGGAGACACAGGCTGGGAAAGTTGGAGGCCCAGAAACCGAAAACCTGGAGGACAGCTGGAG 2637
Qy 2611 AAGATCAGCCCAAGAGACCAAGTGAAGAAATCGGCTGTCTGGAACCTGGAGACAAAGATTG 2670
Db 2638 AAGATCAGCCCAAGAGACCAAGTGAAGAAATCGGCTGTCTGGAACCTGGAGACAAAGATTG 2697
Qy 2671 CGGAGGCTCAGTCTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2730
Db 2698 CGGAGGCTGAGTCTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2757
Qy 2731 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACGCCCTGAGGCTGCAAGCGGCG 2790
Db 2758 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACGCCCTGAGGCTGCAAGCGGCG 2817
Qy 2791 GCCCTGGAGAGCCAGCTTCCGAGCGGAGACAGAGCTGGAAGAGACCAAGAGAGCT 2850
Db 2818 GCCCTGGAGAGCCAGCTTCCGAGCGGAGACAGAGCTGGAAGAGACCAAGAGAGCT 2877
Qy 2851 GAAAGGAGATCCAGGACATCTCAGGACATAGAGATGAATTCAGCGCAAAATTTGATGCT 2910
Db 2878 GAAAGGAGATCCAGGACATCTCAGGACATAGAGATGAATTCAGCGCAAAATTTGATGCT 2937
Qy 2911 CTTCTGAACAGCTGTACTGTAAATCAGAGCTGGAGGAGCAGCTAAACACAGCTGACCGAG 2970
Db 2938 CTTCTGAACAGCTGTACTGTATCAGAGCTGGAGGAGCAGCTAAACACAGCTGACCGAG 2997
Qy 2971 GACAACTGCTGAACCTCAACAAACAAAACTTCTACTTTGTCAAAACTCGATGAGGCTTCT 3030
Db 2998 GACAACTGCTGAACCTCAACAAACAAAACTTCTACTTTGTCAAAACTCGATGAGGCTTCT 3057
Qy 3031 GGCCTCAAGCAGAGATTTGATCACTCGAAGTGAAGTGAACCATCTCCGCGGAGATC 3090
Db 3058 GGCCTCAAGCAGAGATTTGATCACTCGAAGTGAAGTGAACCATCTCCGCGGAGATC 3117
Qy 3091 ACGAAACAGAGATGCGAGCTTACAGCCAGAGCAAAACGATGGAGGCTCTGAAGACACG 3150
Db 3118 ACGAAACAGAGATGCGAGCTTACAGCCAGAGCAAAACGATGGAGGCTCTGAAGACACG 3177
Qy 3151 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAACGATGAGCTGTAGAA 3210
Db 3178 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAACGATGAGCTGTAGAA 3237
Qy 3211 AAAGAGCGGAGTGGGAGGCTGGAGAGGCTCTCGGCTGATGAGAAATCCCAGTTTGA 3270
Db 3238 AAAGAGCGGAGTGGGAGGCTGGAGAGGCTCTCGGCTGATGAGAAATCCCAGTTTGA 3297
Qy 3271 TGTGCGGTTGAGAGCTGAGAGAAATGCTGGAACCCAGAGAAAACAGAGCAGGCGGAGAGCC 3330
Db 3298 TGTGCGGTTGAGAGCTGAGAGAAATGCTGGAACCCAGAGAAAACAGAGCAGGCGGAGAGCC 3357
Qy 3331 GATCAGCGGATCACGAGTCTCGCAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTCAAGGCT 3390
Db 3358 GATCAGCGGATCACGAGTCTCGCAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTCAAGGCT 3417
Qy 3391 GAGATTCGCTGTGAGAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3450
Db 3418 GAGATTCGCTGTGAGAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3477
Qy 3451 GACAAGCTCAATGATCTGGAGAGAGAGCATGCTATGCTTGAATGAATGCTCCCAAGCTTA 3510
Db 3478 GACAAGCTCAATGATCTGGAGAGAGAGCATGCTATGCTTGAATGAATGCTCCCAAGCTTA 3537
Qy 3511 CAGCAGAGCTGGAGACTGAACAGAGCTCAACAGAGGCTTCTGGAAGAGCAGGCAAA 3570
Db 3538 CAGCAGAGCTGGAGACTGAACAGAGCTCAACAGAGGCTTCTGGAAGAGCAGGCAAA 3597

QY	3571	TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGCTCTGACTCAAGGACTGC	3630	Db	4633	GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACCACTCTGTCGCCCGG	4692
Db	3598	TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGCTCTGACTCAAGGACTGC	3657	QY	4654	-----	4653
QY	3631	GAAGCTCTAGATCGGGCTGATCTAATGAAGAAGAAAGTGAATGGAGTATCAGCTG	3690	Db	4693	AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGTCAACCGCTTA	4752
Db	3658	GAAGCTCTAGATCGGGCTGATCTAATGAAGAAGAAAGTGAATGGAGTATCAGCTG	3717	QY	4654	-----	4680
QY	3691	GAAGCATTACAGTCTCTATTCTCATGAAAAAGTGAATGGAGGCACTATTCTCAA	3750	Db	4753	GAATCAGTTGTCGACGCTGGAGAGTTTCTAGGGAAAAAGCAGAAGCTGATGCTAAACTG	4812
Db	3718	GAAGCATTACAGTCTCTATTCTCATGAAAAAGTGAATGGAGGCACTATTCTCAA	3777	QY	4681	CTTGGAAATCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGCAACGCTG	4740
QY	3751	CAAAACAACTCATTTGATTTCTGCAAGCCAAATGGACCAACCTGCTAAAAAGAAAG	3810	Db	4813	CTTGGAAATCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGCAACGCTG	4872
Db	3778	CAAAACAACTCATTTGATTTCTGCAAGCCAAATGGACCAACCTGCTAAAAAGAAAG	3835	QY	4741	CCCTTCAGTGACCAAGTGTGTGCTGGGCACCGAGGAAGGGTCTACGCGCTGGAATGTC	4800
QY	3811	GGTTTATTAGTCGAGGAAAGGAGACCTGCTTTACCCACACAGAGTTCCTCTGAGTAC	3870	Db	4873	CCCTTCAGTGACCAAGTGTGTGCTGGGCACCGAGGAAGGGTCTACGCGCTGGAATGTC	4932
Db	3836	-----	3852	QY	4801	TTGAAAACTCCCTAAACCCATGTCCCAAGGAATGGAGAGTCTTCCAAATTTATATTATC	4860
QY	3871	AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGTGTGCAGAGCTAGAGGAAGCC	3930	Db	4933	TTGAAAACTCCCTAAACCCATGTCCCAAGGAATGGAGAGTCTTCCAAATTTATATTATC	4992
Db	3853	AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGTGTGCAGAGCTAGAGGAAGCC	3912	QY	4861	AAGGACCTGGAGAGTACTCATGATAGCAGGAGAGGCGGACACTGTGCTTGTGGAC	4920
QY	3931	CTTCAGAAACCGGATCAGCTCCGGTCCGCCGGGAGGAGCTGCCACCGCAAGCA	3990	Db	4993	AAGGACCTGGAGAGTACTCATGATAGCAGGAGAGGCGGACACTGTGCTTGTGGAC	5052
Db	3913	CTTCAGAAACCGGATCAGCTCCGGTCCGCCGGGAGGAGCTGCCACCGCAAGCA	3972	QY	4921	GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA	4980
QY	3991	ACGGACCCACACCCATCCAGCCAGCCACCGGAGGAGCAGATGCCATGTCCGCC	4050	Db	5053	GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA	5112
Db	3973	ACGGACCCACACCCATCCAGCCAGCCACCGGAGGAGCAGATGCCATGTCTGCC	4032	QY	4981	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGCGAGCAGATTTGAGAAC	5040
QY	4051	ATGCTCGGTCCGAGAGCAGAGCCAGTGCATGAGCTCTGCGCCCGCCATCCAGC	4110	Db	5113	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGCGAGCAGATTTGAGAAC	5172
Db	4033	ATGCTCGGTCCGAGAGCAGAGCCAGTGCATGAGCTCTGCGCCCGCCATCCAGC	4092	QY	5041	GGGCTCTGCACTCTGAGCAGCATGCCAGCAAAAGTCTGCTATCTCCGCTACACGAAAC	5100
QY	4111	CCGAGAAAGGAGTCTTCACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCAATGC	4170	Db	5173	GGGCTCTGCACTCTGAGCAGCATGCCAGCAAAAGTCTGCTATCTCCGCTACACGAAAC	5232
Db	4093	CCGAGAAAGGAGTCTTCACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCAATGC	4152	QY	5101	CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5160
QY	4171	CACAAATATCTCAGCGATTCACGATAGGACTGAACATCGAGCCACAAAGTGTCTGTG	4230	Db	5233	CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5292
Db	4153	CACAAATATCTCAGCGATTCACGATAGGACTGAACATCGAGCCACAAAGTGTCTGTG	4212	QY	5161	TTCCACCAATACAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG	5220
QY	4231	TGTCTGGATACCGTGCACTTTGGAGCCAGGATCCAAATGTCTCGAATGTGAGTGTG	4290	Db	5293	TTCCACCAATACAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG	5352
Db	4213	TGTCTGGATACCGTGCACTTTGGAGCCAGGATCCAAATGTCTAGAAATGTGAGTGTG	4272	QY	5221	TACAGCTCGAGGAATTCCTGGATGAAGATGAACCAATTCCTTGGCACTGTGTGTTGCC	5280
QY	4291	TGTCACCCCAAGTCTCCAGCTCTGCGAGCCACCTGCGGTGCTGCTGAATATGCC	4350	Db	5353	TACAGCTCGAGGAATTCCTGGATGAAGATGAACCAATTCCTTGGCACTGTGTGTTGCC	5412
Db	4273	TGTCACCCCAAGTCTCCAGCTCTGCGAGCCACCTGCGGTGCTGCTGAATATGCC	4332	QY	5281	GCCTCTTCCACAGCTTCCCTGCTCAATCGTGCAGGTGAACGCGGAGGCGAGGAG	5340
QY	4351	ACACATTCACCGAGGCTTCTGCGGTGAACAAATGAATCCCGAGTCTCCAGACCAAG	4410	Db	5413	GCCTCTTCCACAGCTTCCCTGCTCAATCGTGCAGGTGAACGCGGAGGCGAGGAG	5472
Db	4333	ACACATTCACCGAGGCTTCTGCGGTGAACAAATGAATCCCGAGTCTCCAGACCAAG	4392	QY	5341	GAGTACTTGTGTGTTTCCAGGAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC	5400
QY	4411	GAGCCAGAGCAGCTTGACCTGGAGAGGAGTGAATGAAGTGCACAGGAAATACAAACGA	4470	Db	5473	GAGTACTTGTGTGTTTCCAGGAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC	5532
Db	4393	GAGCCAGAGCAGCTTGACCTGGAGAGGAGTGAATGAAGTGCACAGGAAATACAAACGA	4452	QY	5401	CGCAGAGCAGTCTCAAGTGGAGTCCGCTTACCTTTGGCCCTTGGCTACAGGAACCCCTAT	5460
QY	4471	GGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGATCAAAAGTCTCATTTAT	4530	Db	5533	CGCAGAGCAGTCTCAAGTGGAGTCCGCTTACCTTTGGCCCTTGGCTACAGGAACCCCTAT	5592
Db	4453	GGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGATCAAAAGTCTCATTTAT	4512	QY	5461	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATGAGATCCAGGCACTCTCTCAGCA	5520
QY	4531	GACAAATGAGCCAGAGAGCTGACAGAGCCGGTGGAGAAATTTGAGTGTGCTTCCC	4590	Db	5593	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATGAGATCCAGGCACTCTCTCAGCA	5652
Db	4513	GACAAATGAGCCAGAGAGCTGACAGAGCCGGTGGAGAAATTTGAGTGTGCTTCCC	4572	QY	5521	GGGACCCCTGCGGAGGCTACCTGGAATCCCGAAACCCGGCTACCTGGGCGCTGCCATT	5580
QY	4591	GACGGAGTGTATCTATTATGCTGCGCTTGTGCTTCCGAACCTCCAAATACAGCCAAA	4650	Db	5653	GGGACCCCTGCGGAGGCTACCTGGAATCCCGAAACCCGGCTACCTGGGCGCTGCCATT	5712
Db	4573	GACGGAGTGTATCTATTATGCTGCGCTTGTGCTTCCGAACCTCCAAATACAGCCAAA	4632	QY	5581	TCCTCAGAGAGGATTTACTTGGCGTCTCTCATACAGGATAAATTAAGGTCATTTGCTGC	5640
QY	4651	GCA-----	4653				

Db 5713 TCCTCAGGAGGATTTACTTGGCGTCTCTATACAGGATTAATTAAGGTCATTGCTGC 5772
 Qy 5641 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5700
 Db 5773 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5832
 Qy 5701 AGCAGCCCCACACGAGGAGGCGCCACACCTACACGAGCAGATCACCAGCGCTGGCC 5760
 Db 5833 AGCAGCCCCACACGAGGAGGCGCCACACCTACACGAGCAGATCACCAGCGCTGGCC 5892
 Qy 5761 TCCAGCCCCAGCGCCCGCCAGAGGCCAGCCACCCGCGAGAGCAACACACCCACCGC 5820
 Db 5893 TCCAGCCCCAGCGCCCGCCAGAGGCCAGCCACCCGCGAGAGCAACACACCCACCGC 5952
 Qy 5821 TACCGGAGGGGCGGAGCGAGCTCGCGAGGAGCAAGTCTCTGGCGGCCCGCTGGAGCGA 5880
 Db 5953 TACCGGAGGGGCGGAGCGAGCTCGCGAGGAGCAAGTCTCTGGCGGCCCGCTGGAGCGA 6012
 Qy 5881 GAGAAGTCCCCCGCCGCGATGCTCAGACACCGGAGAGCGGTCCCCCGGAGGCTGTT 5940
 Db 6013 GAGAAGTCCCCCGCCGCGATGCTCAGACACCGGAGAGCGGTCCCCCGGAGGCTGTT 6072
 Qy 5941 GAAGACAGCAGCGGCGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCACAGGTG 6000
 Db 6073 GAAGACAGCAGCGGCGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCACAGGTG 6132
 Qy 6001 AACAGGGAAGAGGCA 6017
 Db 6133 AACAGGTGAGGAGCA 6149

RESULT 11
 ADA05653
 ID ADA05653 standard; cDNA; 6201 BP.
 XX
 AC ADA05653;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human NOV1g encoding cDNA SEQ ID NO:13.
 XX
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..6201
 FT /*tag= a
 FT /*product= "NOV1g"
 XX
 PN WO2003029424-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031373.
 XX
 PR 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328949P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX
 XX (CUPA-) CURAGEN CORP.
 XX
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DM, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 DR WPI; 2003-381626/36.
 DR P-PSDB; ADA05654.
 XX
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 FT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 FT pharmacogenomics.
 XX
 PS Claim 20; Page 104-105; 586pp; English.
 XX
 CC The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity of or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence encodes a human NOVX protein from
 CC the present invention.
 XX

SQ Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;
Query Match 85.6%; Score 5629.8; DB 7; Length 6201;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;
QY 19 ATGTTGAAGTTCACAAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT 78
Db 1 ATGTTGAAGTTCACAAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT 60
QY 79 GCACGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACCAACCTTTATGACTCAA 138
Db 61 GCACGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACCAACCTTTATGACTCAA 120
QY 139 CAGCAGATGTCCTCTTTTCCCGAGAAGGATATTTAGATGCCCTCTTTGTTCTCTTTGAA 198
Db 121 CAGCAGATGTCCTCTTTTCCCGAGAAGGATATTTAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 199 GAATGAGTCAGCTGCTCTGATGAGATTAAGCACTGAGCAACTTTGTCGGAAGTAT 258
Db 181 GAATGAGTCAGCTGCTCTGATGAGATTAAGCACTGAGCAACTTTGTCGGAAGTAT 240
QY 259 TCCGACACATAGCTGAGTTACAGAGCTCCAGCTTCCGCAAGGACTTGAAGTCAGA 318
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCTTCCGCAAGGACTTGAAGTCAGA 300
QY 319 AGTCTGTAGTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAGAAACCGGG 378
Db 301 AGTCTGTAGTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAGAAACCGGG 360
QY 379 GACATCTATGCTATGAAAGTATGAGAGAGGCTTTATTTGGCCAGGAGCAGGTTTCA 438
Db 361 GACATCTATGCTATGAAAGTATGAGAGAGGCTTTATTTGGCCAGGAGCAGGTTTCA 420
QY 439 TTTTGTGAGGAAGCGGAACATATTTATCTGAAGCACAAGCCCGTGGATCCCCAATTA 498
Db 421 TTTTGTGAGGAAGCGGAACATATTTATCTGAAGCACAAGCCCGTGGATCCCCAATTA 480
QY 499 CAGTATGCTTTGAGCAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 558
Db 481 CAGTATGCTTTGAGCAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
QY 559 GACTTGTCTCACTTTTGAATAGATGAGGACCACTTATGATGAAACCTGATACAGTTT 618
Db 541 GACTTGTCTCACTTTTGAATAGATGAGGACCACTTATGATGAAACCTGATACAGTTT 600
QY 619 TACCTAGCTGAGCTGATTTGGCTGTTACACGCTTCATCTGATGGGATAGTGATCGA 678
Db 601 TACCTAGCTGAGCTGATTTGGCTGTTACACGCTTCATCTGATGGGATAGTGATCGG 660
QY 679 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 738
Db 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 720
QY 739 GGATCTGCCCGGAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 798
Db 721 GGATCTGCCCGGAAATGAATTCAAACA--GGTGAATGCCAAACTCCCGATTGGGACC 777
QY 799 CAGATTAATAGCTCCTGAACTGCTGATGATGAACGGGGATGGAAAGGACCTAC 858
Db 778 CAGATTAATAGCTCCTGAACTGCTGATGATGAACGGGGATGGAAAGGACCTAC 837
QY 859 GGCTCGGACTGTGACTGGTGGTCAAGTGGGCTGATTGCCCTATGAGATGATTTATGGGAGA 918
Db 838 GGCTCGGACTGTGACTGGTGGTCAAGTGGGCTGATTGCCCTATGAGATGATTTATGGGAGA 897
QY 919 TCCCCCTTCGAGAGGAAACCTCTGCCAGAACCTTCAATTAACATTAATGAAATTCAGCGG 978
Db 898 TCCCCCTTCGAGAGGAAACCTCTGCCAGAACCTTCAATTAACATTAATGAAATTCAGCGG 957
QY 979 TTTTTCGAATTTCCAGATGACCCCAAGTGAGCAGTACTTTCTTCTGATCTGATTCAGAGC 1038
Db 958 TTTTTCGAATTTCCAGATGACCCCAAGTGAGCAGTACTTTCTTCTGATCTGATTCAGAGC 1017

QY 1039 TTGTTGTCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1098
Db 1018 TTGTTGTCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1077
QY 1099 TCTAAATTTGACTGGAACAACATTTGTAATCTCTCTCCCCCCTTCTGTTCCCAACCTCAAG 1158
Db 1078 TCTAAATTTGACTGGAACAACATTTGTAACGCTCTCTCCCCCCTTCTGTTCCCAACCTCAAG 1137
QY 1159 TCTGACCATCACACCTCCAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTTCATCTCT 1218
Db 1138 TCTGACCATCACACCTCCAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTTCATCTCT 1197
QY 1219 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAAGAACTGCGTTTGTGGGTTTTCG 1278
Db 1198 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAAGAACTGCGTTTGTGGGTTTTCG 1257
QY 1279 TACAGCAAGCACTGGGATTTCTTGTAGATCTGAGTCTGTTCTGTCGGGCTCGGACTCC 1338
Db 1258 TACAGCAAGCACTGGGATTTCTTGTAGATCTGAGTCTGTTCTGTCGGGCTCGGACTCC 1317
QY 1339 CTTGCCAAGACTAGCTCCATCGAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1398
Db 1318 CTTGCCAAGACTAGCTCCATCGAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1377
QY 1399 TCTCAGACAAGTGTCAAAAGTGGAGCAAGAAATGACCCGTTACATCGGAGGTGCA 1458
Db 1378 TCTCAGACAAGTGTCAAAAGTGGAGCAAGAAATGACCCGTTACATCGGAGGTGCA 1437
QY 1459 GAGGTGAGGCTGTGCTTAGTCAAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518
Db 1438 GAGGTGAGGCTGTGCTTAGTCAAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497
QY 1519 TCCCTCTGAGCAGGACCTTGTCTACATACAGAAATCAAGAGCAGGAGTACAGGCTCAA 1578
Db 1498 TCCCTCTGAGCAGGACCTTGTCTACCTACATCACAGAATGCAGTAGCTTAAAGCGAAGT 1557
QY 1579 TTGAGCAAGCAGCGATGGAGGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCCAT 1638
Db 1558 TTGAGCAAGCAGCGATGGAGGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCCAT 1617
QY 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1698
Db 1618 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677
QY 1699 GTGGAAGAAATGAGGTTGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1758
Db 1678 GTGGAAGAAATGAGGTTGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1737
QY 1759 CCGAGTCTATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTCAG 1818
Db 1738 CCGAGTCTATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTCAG 1797
QY 1819 CGGAAAGCAGCAGAAATCTCAGCTAATACTGTTGAAGCTAAGGATCAAGGGAAGCCTGAA 1878
Db 1798 CGGAAAGCAGCAGAAATCTCAGCTAATACTGTTGAAGCTAAGGATCAAGGGAAGCCTGAA 1857
QY 1879 GTGGGAAATATCGAACTGGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938
Db 1858 GTGGGAAATATCGAACTGGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1917
QY 1939 CTCCAAGAGAACTGGGAAGGCT----- 1962
Db 1918 CTCCAAGAGAACTGGGAAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTCGAG 1977
QY 1963 -----GCAAGAGGAGCGAGCGAGGGAGCTGAGAGCTGAGAGCTGCAGAACCGAGAG 2010
Db 1978 AATATCCCGAGGCAAGGAGCGAGGAGCTGGAGAACTGAGAACTGCAGAACCGAGAG 2037
QY 2011 GATTCTTTCTAAGGCATCAGAAAGAACTGGTGGAACTGAGGAACCGCCCAATTTCTG 2070
Db 2038 GATTCTTTCTGAGGCATCAGAAAGAACTGGTGGAACTGAGGAACCGCCCAATTTCTG 2097

4213	TGCTGTGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTAGAAATGTCAGGTGATG	4272
4291	TGTCACCCCAAGTGTCCACGCTCTTCCAGCCACACTGGGGCTTGCTGCTGAATATGCC	4350
4273	TGTCACCCCAAGTGTCCACGCTCTTGGCCAGCCACCTGGGCTTGCTGCTGAATATGCC	4332
4351	ACACACTTCAACGAGGCTTCTCCGCTGACAAATGAATCTCCACGGTCTCCAGACCAAG	4410
4333	ACACACTTCAACGAGGCTTCTCCGCTGACAAATGAATCTCCACGGTCTCCAGACCAAG	4392
4411	GAGCCAGCAGCAGCTTGACCTTGGAGGGTGGATGAAGGTGCCAGGAATACAAACGA	4470
4393	GAGCCAGCAGCAGCTTGACCTTGGAGGGTGGATGAAGGTGCCAGGAATACAAACGA	4452
4471	GGACAGCAAGGCTGGGACAGGAATACATTTGCTCTGGAGGGATCAAAAGTCTCATTTAT	4530
4453	GGACAGCAAGGCTGGGACAGGAATACATTTGCTCTGGAGGGATCAAAAGTCTCATTTAT	4512
4531	GACAATGAAGCAGAGAAGCTGCACAGAGCCGGTGGAGAAATTCAGAGTGTGCCCTTCC	4590
4513	GACAATGAAGCAGAGAAGCTGCACAGAGCCGGTGGAGAAATTCAGAGTGTGCCCTTCC	4572
4591	GACGGGATGTATCTATTCTATGCTGCGGTGGTCTCGAATCGCAATACAGCCAAA	4650
4573	GACGGGATGTATCTATTCTATGCTGCGGTGGTCTCGAATCGCAATACAGCCAAA	4632
4651	GCA-----	4653
4633	GCAGATGTCCCATACATACCTAGAGATGGAATCTACCCGCACACCACTGCTGGCCGGG	4692
4654	-----	4653
4693	AGAACCTCTACTTGCTAGCTCCAGCTTCCCTGCACAAACAGCGCTGGGTACCGGCTTA	4752
4654	-----	4680
4753	GAATCAGTTGTGCAAGTGGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGCTAAACTG	4812
4681	CTTGGAACTCCCTGCTGAACTGGAAAGTGATGACCGTCTAGACATGAACTGCACGCTG	4740
4813	CTTGGAACTCCCTGCTGAACTGGAAAGTGATGACCGTCTAGACATGAACTGCACGCTG	4872
4741	CCCTTCAGTGACCAAGTGGTGTGGTGGGCAACCGAGGAAGGGCTCTACGCCCTGAATGTC	4800
4873	CCCTTCAGTGACCAAGTGGTGTGGTGGGCAACCGAGGAAGGGCTCTACGCCCTGAATGTC	4932
4801	TTGAAAACCTCCCTAACCCATGTCACCAAGTGGAGCAGTCTTCCAAATTTATATTATC	4860
4933	TTGAAAACCTCCCTAACCCATGTCACCAAGTGGAGCAGTCTTCCAAATTTATATTATC	4992
4861	AAGACCTGGAGAGCTACTCATGATAGCAGGAAGCGGCACTGTGCTCTTGCGAC	4920
4993	AAGACCTGGAGAGCTACTCATGATAGCAGGTGAAGAGCGGCACTGTGCTCTTGCGAC	5052
4921	GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA	4980
5053	GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA	5112
4981	CCCAACATTTTGAAGCTGTCAAGGGCTGCGACCTGTTTGGGGCAGGCAAGATTGAGAAC	5040
5113	CCCAACATTTTGAAGCTGTCAAGGGCTGCGACCTGTTTGGGGCAGGCAAGATTGAGAAC	5172
5041	GGGCTCTGCATCTGTGAGCCATCCAGCAAGTGTCTATTCTCCGCTACACGAAAC	5100
5173	GGGCTCTGCATCTGTGAGCCATCCAGCAAGTGTCTATTCTCCGCTACACGAAAC	5232
5101	CTCAGCAAAATACTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5160
5233	CTCAGCAAAATACTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5292
5161	TTACCAATTTACGTATCTCTATTGGAAACCAATTAATTTCTACGAATTCGATGAACGAG	5220
5293	TTACCAATTTACGTATCTCTATTGGAAACCAATTAATTTCTACGAATTCGATGAACGAG	5352

RESULT 12
AB078871

ABQ78871
ID ABQ78871 standard; cDNA; 5877 BP.

XX
AC

DT 10-OCT-2002 (first entry)

DI 10 OCT 2002 12:13:33
XX
DE Human kinase cDNA #2.

Human: Kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;
XX
KW Human; enzyme; serine-threonine kinase; nootropic; cytostatic;
KW Cytroin rho-interacting kinase; gene therapy; mental disorder; cancer;
KW gene; ss.

XX
OS

XX Key Location/Qualifiers
FH 1..5877
PT /*tag= a
FT /product= "Kinase"
XX WO200259325-A2.
XX 01-AUG-2002.
XX 20-DEC-2001; 2001WO-US050497.
XX 27-DEC-2000; 2000US-0258335P.
XX (LEXI-) LEXICON GENETICS INC.
XX Yu X, Miranda M, Friddle CJ;
XX WPI; 2002-599796/64.
XX P-PSDB; ABB81928.
XX Novel polynucleotide encoding human proteins that are structurally
PT similar to animal kinases, useful for drug screening, diagnosis, in gene
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
PT applications.
XX Disclosure; Page 44-45; 50pp; English.
XX
XX The invention relates to a novel human protein that shares structural
CC similarity with animal kinases, including serine-threonine kinases,
CC particularly Citron rho-interacting kinases. The proteins of the
CC invention have nootropic and cytosolic activity. The polynucleotides may
CC have a use in gene therapy. The encoded novel polypeptides are useful for
CC generating antibodies, as reagents in diagnostic assays, for identifying
CC other cellular gene products related to NHP and as reagents in assays for
CC screening for compounds that are useful in the treatment of mental,
CC biological or medical disorders and diseases including cancer. The
CC sequence encodes a novel human kinase of the invention
XX
XX Sequence 5877 BP; 1680 A; 1433 C; 1583 G; 1181 T; 0 U; 0 Other;
SQ
Query Match 81.5%; Score 5358; DB 6; Length 5877;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5634; Conservative 0; Mismatches 5; Indels 243; Gaps 3;
QY 19 ATGTTGAAGTTCAATATGAGCGCGGAAATCTTTGGATGCTGGTCTGAAACCCATT 78
DB 1 ATGTTGAAGTTCAATATGAGCGCGGAAATCTTTGGATGCTGGTCTGAAACCCATT 60
QY 79 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 138
DB 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 120
QY 139 CAGCAGATGTCCTCTTTCCGAGAAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198
DB 121 CAGCAGATGTCCTCTTTCCGAGAAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 199 GAATGAGTCAGCTCTCTGATGAAGATTAAAGCAGTCGAGCAACTTTGTCGGGAAGTAT 258
DB 181 GAATGAGTCAGCTCTCTGATGAAGATTAAAGCAGTCGAGCAACTTTGTCGGGAAGTAT 240
QY 259 TCGACACCATAGCTCAGTTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAGTCAGA 318
DB 241 TCGACACCATAGCTCAGTTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAGTCAGA 300
QY 319 AGTCTGTAGTGTGGTCACTTTCGAGTGAAGTCGAGTCGAGTGAAGAGGAGGAGGAGG 378
DB 301 AGTCTGTAGTGTGGTCACTTTCGAGTGAAGTCGAGTCGAGTGAAGAGGAGGAGGAGG 360
QY 379 GACATCTATGCTATGAAGTGTAGAGAGAGAGGCTTTATTTGGCCCGAGGAGGAGGTTCA 438
DB 361 GACATCTATGCTATGAAGTGTAGAGAGAGGCTTTATTTGGCCCGAGGAGGAGGTTCA 420

QY 439 TTTTTCAGGAGAGCGGAAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 498
DB TTTTTCAGGAGAGCGGAAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480
QY 499 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 558
DB CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540
QY 559 GACTTGCTGTCACTTTTGAATAGATATGAGACACAGTTAGATGAAAAACCTGTACAGTTT 618
DB GACTTGCTGTCACTTTTGAATAGATATGAGACACAGTTAGATGAAAAACCTGTACAGTTT 600
QY 619 TACCTAGCTAGCTGATTTTGGCTTTCACAGGCTTCAITCTGATGGATACGTGATCGA 678
DB TACCTAGCTAGCTGATTTTGGCTTTCACAGGCTTTCATCTGATGGATACGTGATCGA 660
QY 679 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGATTT 738
DB GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGATTT 720
QY 739 GGAATCTGCGCGAAAAATGAATTTCAACAAGATGGTGAATCCCAACCTCCGATTTGGGACC 798
DB GGAATCTGCGCGAAAAATGAATTTCAACAAGATGGTGAATCCCAACCTCCGATTTGGGACC 780
QY 799 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGACGGGATGAAAAAGGACCTAC 858
DB CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGACGGGATGAAAAAGGACCTAC 840
QY 859 GGCCTGACCTGTGACTCGTGTGCTAGTGGGCTGATTCCTATGAGATGATTTATGGAGA 918
DB GGCCTGACCTGTGACTCGTGTGCTAGTGGGCTGATTCCTATGAGATGATTTATGGAGA 900
QY 919 TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 978
DB TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 960
QY 979 TTTTGAATTTCCAGATGACCCCAAGTCAAGTCAAGTCTTCTGATCTGATTTCAAGC 1038
DB TTTTGAATTTCCAGATGACCCCAAGTCAAGTCAAGTCTTCTGATCTGATTTCAAGC 1020
QY 1039 TTGTTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1098
DB TTGTTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1080
QY 1099 TCTAAATTCAGCTGGACACATTCGTAACCTCTCCCTCCCTTCCTCCACCTCAAG 1158
DB TCTAAATTCAGCTGGACACATTCGTAACCTCTCCCTCCCTTCCTCCACCTCAAG 1140
QY 1159 TCTGACGATGACACCTCCAAATTTTGAATGAACACAGAGAAGAAATTCGTGGTTCATCTCT 1218
DB TCTGACGATGACACCTCCAAATTTTGAATGAACACAGAGAAGAAATTCGTGGTTCATCTCT 1200
QY 1219 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTCCGTTTGGGGTTTCG 1278
DB CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTCCGTTTGGGGTTTCG 1260
QY 1279 TACAGCAAGCACTGGGGATTTGTTGATGATCTGAGTCTGTTGCTGGGCTCTGGACTCC 1338
DB TACAGCAAGCACTGGGGATTTGTTGATGATCTGAGTCTGTTGCTGGGCTCTGGACTCC 1320
QY 1339 CTGCGCAAGCACTAGCTCCATGAAAAAGAACTTCTCATCAAAAAGCAAGAGCTTCAAGAC 1398
DB CTGCGCAAGCACTAGCTCCATGAAAAAGAACTTCTCATCAAAAAGCAAGAGCTTCAAGAC 1380
QY 1399 TCTCAGGACAAAGTGTCAACAGATGGAGCAGAAATGACCCGTTTACATCGGAGAGTCA 1458
DB TCTCAGGACAAAGTGTCAACAGATGGAGCAGAAATGACCCGTTTACATCGGAGAGTCA 1440
QY 1459 GAGTGGAGGCTGTGCTTATGTCAGAGGAGGTGGAAGTGAAGCCCTCTGAGACTCAGAGA 1518
DB GAGTGGAGGCTGTGCTTATGTCAGAGGAGGTGGAAGTGAAGCCCTCTGAGACTCAGAGA 1500
QY 1519 TCCCTCTCGAGCAGGACCTTGCTACCTACATCACAGATGCAAGTCTTAAAGCAAGT 1578

Db | 1501 | TCCCTCCCTGGAGCAGGACCTTCTACCTACATCAACAGAAATGCAGTAGCTTAAAGCGAAGT | 1560 |
Qy | 1579 | TTGGAGCAAGCACGGATGGAGGTGTCACAGGAGGATGACAAAGCACTGCAGCTTCTCCAT | 1638 |
Db | 1561 | TTGGAGCAAGCACGGATGGAGGTGTCACAGGAGGATGACAAAGCACTGCAGCTTCTCCAT | 1620 |
Qy | 1639 | GATATCAGAGACAGACCGGAGGATCCAGAAATCAAGAGCAGGATACAGGCTCAA | 1698 |
Db | 1621 | GATATCAGAGACAGACCGGAGGATCCAGAAATCAAGAGCAGGATACAGGCTCAA | 1680 |
Qy | 1699 | GTGGAAGAAATGAGGTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA | 1758 |
Db | 1681 | GTGGAAGAAATGAGGTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA | 1740 |
Qy | 1759 | CGGAGTGATCTTACGAATCTGAGCTGAGAGGTCTCGGCTTGCTGCTGAGAAATTCAG | 1818 |
Db | 1741 | CGGAGTGATCTTACGAATCTGAGCTGAGAGGTCTCGGCTTGCTGCTGAGAAATTCAG | 1800 |
Qy | 1819 | CGGAAAGCGACAGATGTCAGCATAACTGTTGGAAGGCTTAGGATCAAGGGAAGCTGAA | 1878 |
Db | 1801 | CGGAAAGCGACAGATGTCAGCATAACTGTTGGAAGGCTTAGGATCAAGGGAAGCTGAA | 1860 |
Qy | 1879 | GTGGAGAAATATGCGAAATCGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG | 1938 |
Db | 1861 | GTGGAGAAATATGCGAAATCGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG | 1920 |
Qy | 1939 | CTCGAAGAAATCGAGAGGCT----- | 1962 |
Db | 1921 | CTCGAAGAAATCGAGAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTGCAG | 1980 |
Qy | 1963 | -----GCAAAGAGCGAGCCGAGAGGAGCTGGAGAAGCTGCGAGAACCGAGAG | 2010 |
Db | 1981 | AATATCCGCGCAGGCAAGAGCGAGCGAGAGGAGCTGGAGAAGCTGCAGAACCGAGAG | 2040 |
Qy | 2011 | GATTCCTCTGGAAGGATCAGAAAGAGCTGTGGAAGCTGAGGAACCGCGGCTTCTCTG | 2070 |
Db | 2041 | GATTCCTCTGGAAGGATCAGAAAGAGCTGTGGAAGCTGAGGAACCGCGGCTTCTCTG | 2100 |
Qy | 2071 | GAGAACAGGTAAAGAGACTAGAGACCATGGAAGCTTAGAGAAACAGACTGAAGATGAC | 2130 |
Db | 2101 | GAGAACAGGTAAAGAGACTAGAGACCATGGAAGCTTAGAGAAACAGACTGAAGATGAC | 2160 |
Qy | 2131 | ATCCAGACAAAATCCCAACAGATCCAGCAGATGCGCTGATAAAATTCGAGCTCGAAGAG | 2190 |
Db | 2161 | ATCCAGACAAAATCCCAACAGATCCAGCAGATGCGCTGATAAAATTCGAGCTCGAAGAG | 2220 |
Qy | 2191 | AAACATCGGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAGTGCCACTGAAACAGAAAGAG | 2250 |
Db | 2221 | AAACATCGGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAGTGCCACTGAAACAGAAAGAG | 2280 |
Qy | 2251 | CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAAGACCTGCTGCAC | 2310 |
Db | 2281 | CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAAGACCTGCTGCAC | 2340 |
Qy | 2311 | AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAGGGGAAA | 2370 |
Db | 2341 | AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAGGGGAAA | 2400 |
Qy | 2371 | ATTCTCAGCGAAACAGAGGCGATGATCAATGATGATGATTCCAGATCAGATCCCTGGAA | 2430 |
Db | 2401 | ATTCTCAGCGAAACAGAGGCGATGATCAATGATGATGATTCCAGATCAGATCCCTGGAA | 2460 |
Qy | 2431 | CAGAGATTGTGGAACTGTCTGAAGCCAAATAAATTCAGAGAAAATAGCAGTCTTTTACC | 2490 |
Db | 2461 | CAGAGATTGTGGAACTGTCTGAAGCCAAATAAATTCAGAGAAAATAGCAGTCTTTTACC | 2520 |
Qy | 2491 | CAAAGAAATAGAGCCCAAGAGAGATGATTTCTGAATCAGGCTCAGCAAGAAATTTTAC | 2550 |
Db | 2521 | CAAAGAAATAGAGCCCAAGAGAGATGATTTCTGAATCAGGCTCAGCAAGAAATTTTAC | 2580 |
Qy | 2551 | CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAAAACTGGAGAGCAGCTGGAG | 2610 |

Db | 2581 | CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAAAACTGGAGGACAGCTGGAG | 2640 |
Qy | 2611 | AAGATCAGCCACCAAGACCACTGACAAAGATCGGCTGCTGGAACTGGAGACAAAGATTG | 2670 |
Db | 2641 | AAGATCAGCCACCAAGACCACTGACAAAGATCGGCTGCTGGAACTGGAGACAAAGATTG | 2700 |
Qy | 2671 | CGGAGGTCAAGTCTAGAGCACGAGGACGAAACTGGAGCTCAAGGCCAGCTCACAGAG | 2730 |
Db | 2701 | CGGAGGTCAAGTCTAGAGCACGAGGACGAAACTGGAGCTCAAGGCCAGCTCACAGAG | 2760 |
Qy | 2731 | CTACAGCTCTCCCTGCGAGGCGCGATCACAGTTGACAGCCTGCGAGGCTGCAOCCGGCG | 2790 |
Db | 2761 | CTACAGCTCTCCCTGCGAGGCGCGATCACAGTTGACAGCCTGCGAGGCTGCAOCCGGCG | 2820 |
Qy | 2791 | GCCTGCGAGGCGCAGCTTCGCCAGGCGAGACAGAGCTGGAAGAGACCAAGCAGAGAGCT | 2850 |
Db | 2821 | GCCTGCGAGGCGCAGCTTCGCCAGGCGAGACAGAGCTGGAAGAGACCAAGCAGAGAGCT | 2880 |
Qy | 2851 | GAAGAGAGATCCAGGCACCTCAGGCGACATAGAGATGAATCCAGGCGCAAAATTTGATGCT | 2910 |
Db | 2881 | GAAGAGAGATCCAGGCACCTCAGGCGACATAGAGATGAATCCAGGCGCAAAATTTGATGCT | 2940 |
Qy | 2911 | CTTGCTAACAGCTGTACTGTAAATCAGAGCTCGAGGAGAGAGCTAAACAGCTGACCGGAG | 2970 |
Db | 2941 | CTTGCTAACAGCTGTACTGTAAATCAGAGCTCGAGGAGAGAGCTAAACAGCTGACCGGAG | 3000 |
Qy | 2971 | GACAACTCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT | 3030 |
Db | 3001 | GACAACTCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT | 3060 |
Qy | 3031 | GCGCCCAACGACAGATTTGTACAACTGCGAAGTGAAGTGAACCATCTCCGCCGGGAGATC | 3090 |
Db | 3061 | GCGCCCAACGACAGATTTGTACAACTGCGAAGTGAAGTGAACCATCTCCGCCGGGAGATC | 3120 |
Qy | 3091 | ACGGACGAGAGATGCGAGCTTACCGCCAGAACCAACGATGGAGGCTCTGAGAGACCAAG | 3150 |
Db | 3121 | ACGGACGAGAGATGCGAGCTTACCGCCAGAACCAACGATGGAGGCTCTGAGAGACCAAG | 3180 |
Qy | 3151 | TGCACCATGCTGGAGGAAACAGGTCATGGAATTTGGAGGCCCTAAACGATGAGCTGTAGAA | 3210 |
Db | 3181 | TGCACCATGCTGGAGGAAACAGGTCATGGAATTTGGAGGCCCTAAACGATGAGCTGTAGAA | 3240 |
Qy | 3211 | AAAGAGCGGAGTGGGAGGCTCGAGAGGCTCTCGGTTGATGAGAAATCCGAGTTTGGAG | 3270 |
Db | 3241 | AAAGAGCGGAGTGGGAGGCTCGAGAGGCTCTCGGTTGATGAGAAATCCGAGTTTGGAG | 3300 |
Qy | 3271 | TGTCCGCTTCGAGAGCTGCGAGAGATGCTGGAACCGAGAAACAGAGCGGCGAGAGCC | 3330 |
Db | 3301 | TGTCCGCTTCGAGAGCTGCGAGAGATGCTGGAACCGAGAAACAGAGCGGCGAGAGCC | 3360 |
Qy | 3331 | GATCAGCGGATCACCGAGTCTCGCAGGTGGTGGAGCTGGCAGTGAAGAGACACAGGCT | 3390 |
Db | 3361 | GATCAGCGGATCACCGAGTCTCGCAGGTGGTGGAGCTGGCAGTGAAGAGACACAGGCT | 3420 |
Qy | 3391 | GAGATTCCTCGCTCGCAGGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCTCTCT | 3450 |
Db | 3421 | GAGATTCCTCGCTCGCAGGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCTCTCT | 3480 |
Qy | 3451 | GACAAGCTCAATGACCTTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAGAGCTTA | 3510 |
Db | 3481 | GACAAGCTCAATGACCTTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAGAGCTTA | 3540 |
Qy | 3511 | CAGCAGAGCTGAGATGAAACAGAGCTCAACAGAGGCTTCTGAGAGAGAGCCAAA | 3570 |
Db | 3541 | CAGCAGAGCTGAGATGAAACAGAGCTCAACAGAGGCTTCTGAGAGAGAGCCAAA | 3600 |
Qy | 3571 | TTACAGCAGAGATGGAACCTGCAAAAAATCACAATTTCCGCTCTGACTCAAGGACTGCAA | 3630 |
Db | 3601 | TTACAGCAGAGATGGAACCTGCAAAAAATCACAATTTCCGCTCTGACTCAAGGACTGCAA | 3660 |
Qy | 3631 | GAAGCTTAGATCGGCTGATCTACTGAAAGACAGAAAGATGACTTGGAGTATCAGCTG | 3690 |
Db | 3661 | GAAGCTTAGATCGGCTGATCTACTGAAAGACAGAAAGATGACTTGGAGTATCAGCTG | 3720 |

CC	cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease, to
CC	enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX	
SQ	Sequence 6609 Bp; 1694 A; 1779 C; 1646 G; 1490 T; 0 U; 0 Other;
	Query Match 52.2%; Score 3475.2; DB 3; Length 6609;
	Best Local Similarity 94.2%; Pred. No. 0;
	Matches 3710; Conservative 0; Mismatches 33; Indels 195; Gaps 2;
QY	2273 AAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAATGA 2332
DB	299 AGGTGTTGGACAAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAATGA 358
QY	2333 TGCAGAGACACGAGGAGGAGGCCCATGAGNAGGGCAAAATTTCTCAGGAAACAGAGGCGA 2392
DB	359 TGCAGAGACACGAGGAGGAGGCCCATGAGNAGGGCAAAATTTCTCAGGAAACAGAGGCGA 418
QY	2393 TGAATCAATGCTATGGGATTTCCAAGATCAGATCCCTGGAAACAGAGGATTTGTGGAACTGTCTG 2452
DB	419 TGAATCAATGCTATGGATTTCCNAGATCAGATCCCTGGAAACAGAGGATTTGTGGAACTGTCTG 478
QY	2453 AAGCCAAATAAATTTCGAGCAATATAGAGTCTTTTACCCAAAGGAACATGAAGGCCCAAG 2512
DB	479 AAGCCAAATAAATTTCGAGCAATATAGAGTCTTTTACCCAAAGGAACATGAAGGCCCAAG 538
QY	2513 AAGAGATGATTTCTGAACTCAGGCAACAGAGAAATTTTACCTGGAGACACAGGCTGGGAAGT 2572
DB	539 AAGAGATGATTTCTGAACTCAGGCAACAGAGAAATTTTACCTGGAGACACAGGCTGGGAAGT 598
QY	2573 TGGAGGCCAGAAACCGAAAACTCGAGAGCAGCTGGAGAAAGATCAGCCACAGAACCA 2632
DB	599 TGGAGGCCAGAAACCGAAAACTCGAGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACA 658
QY	2633 GTGACAAGAAATCGGCTGCTGGAACCTGGAGACAGAGATTGCGGAGGTCAGTCTAGAGCAAG 2692
DB	659 GTGACAAGAAATCGGCTGCTGGAACCTGGAGACAGAGATTGCGGAGGTCAGTCTAGAGCAAG 718
QY	2693 AGGAGCAGAAACTGAGAGCTCAAGCGCCAGCTCA CAGAGCTACAGCTCTCCCTCGAGAGC 2752
DB	719 AGGAGCAGAAACTGAGAGCTCAAGCGCCAGCTCA CAGAGCTACAGCTCTCCCTCGAGAGC 778
QY	2753 GCAGTCA CAGTTGACAGCCCTGCAGGCTGCAGCGGCGGCCCTGGAGAGCCAGCTTCGCC 2812
DB	779 GCAGTCA CAGTTGACAGCCCTGCAGGCTGCAGCGGCGGCCCTGGAGAGCCAGCTTCGCC 838
QY	2813 AGGCGAAGACAGAGCTGGAAAGAGACCA CAGCAGAAGCTGAAGAGGAGATCCAGGCACTCA 2872
DB	839 AGGCGAAGACAGAGCTGGAAAGAGACCA CAGCAGAAGCTGAAGAGGAGATCCAGGCACTCA 898
QY	2873 CGGCA CATA GAGATGA AATCCAGCGCA AATTTGATGCTCTTTCGTAA CAGCTGTACTGTAA 2932
DB	899 CGGCA CATA GAGATGA AATCCAGCGCA AATTTGATGCTCTTTCGTAA CAGCTGTACTGTAA 958
QY	2933 TCACAGACCTTGGAGGAGCAGCTAAAACAGCTGACGAGGACCAACGCTGAACTCAAACAAC 2992
DB	959 TCACAGACCTTGGAGGAGCAGCTAAAACAGCTGACGAGGACCAACGCTGAACTCAAACAAC 1018
QY	2993 AAAA ACTTCA TTTGTCCAAA CAACTCGATGAGGCTTTTGGCGCCAA CAGCAGATTTGTATC 3052
DB	1019 AAAA ACTTCA TTTGTCCAAA CAACTCGATGAGGCTTTTGGCGCCAA CAGCAGATTTGTATC 1078
QY	3053 AACTCGGAAGTGAAGTGGACCACTTCGCGCGGAGATCAGCGAA CAGAGATCAGCTTA 3112
DB	1079 AACTCGGAAGTGAAGTGGACCACTTCGCGCGGAGATCAGCGAA CAGAGATCAGCTTA 1138
QY	3113 CCAGCCAGAGCAAAACGATGAGGCTCTGAAGACCA CAGTCACCACTGCTGGAGGAAACAGG 3172
DB	1139 CCAGCCAGAGCAAAACGATGAGGCTCTGAAGACCA CAGTCACCACTGCTGGAGGAAACAGG 1198
QY	3173 TCATGGA TTTGGAGGCCCTTAAACGATGAGCTGCTAGAAAAG CCGCAGTGGAGGCGCT 3232
DB	1199 TCATGGA TTTGGAGGCCCTTAAACGATGAGCTGCTAGAAAAG CCGCAGTGGAGGCGCT 1258

Db 5836 AG 5837

RESULT 13
AAC77568
ID AAC77568 standard; cDNA; 6609 BP.
XX AC AAC77568;
XX XX
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF3123 polynucleotide sequence SEQ ID NO:6245.
XX XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnary; antipsoriatic; antiParkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antihydroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disease; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.

XX XX
OS Homo sapiens.
PN WC200058473-A2.
XX XX
PD 05-OCT-2000.
XX XX
PF 31-MAR-2000; 2000WO-US008621.
XX PR 31-MAR-1999; 99US-01276507P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX PA
PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX XX
DR WPI: 2000-602362/57.
DR P-PDSB; AAB43359.
XX XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX XX
PS Claim 5; Page 5429-5433; 5507pp; English.
XX XX
CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipsoriatic; antiParkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antihydroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC CC

Db	3374	ATAGAATGACCAATTCCTTTGGCACTGCTGTGTTTGGCGCTCTTCCAAAGCTTCCCTG	3433
Qy	5303	TCTCAATCGTGCAGGTGAACAGCGCCAGGGCAGGAGAGAGTAGTACTTGTCTGTGTTTCCACG	5362
Db	3434	TCTCAATCGTGCAGGTGAACAGCGCCAGGGCAGGAGAGAGTAGTACTTGTCTGTGTTTCCACG	3493
Qy	5363	AATTGGAGTGTTCTGTGGATTCTTTACGGAAGACGTAGCCGACAGAGCATCTCAAGTGGGA	5422
Db	3494	AATTGGAGTGTTCTGTGGATTCTTTACGGAAGAGGTAGCCGACAGAGCATCTCAAGTGGGA	3553
Qy	5423	GTCCGCTTACCTTTGGCCTTTGCCCTACAGAGAACCCCTATCTGTTTGTGACCCACTTCAACT	5482
Db	3554	GTCCGCTTACCTTTGGCCTTTGCCCTACAGAGAACCCCTATCTGTTTGTGACCCACTTCAACT	3613
Qy	5483	CACCTCGAAGTAATTGAGATCAGGACAGCTCTCTCAGCAGGGAACCCCTGCCGAGCGTACC	5542
Db	3614	CACCTCGAAGTAATTGAGATCAGGACAGCTCTCTCAGCAGGGAACCCCTGCCGAGCGTACC	3673
Qy	5543	TGACATCCCGAACCCTGGCGCTACCTGGGCCCTTGGCATTTCCTCAGGAGCGAATTACTTGG	5602
Db	3674	TGACATCCCGAACCCTGGCGCTACCTGGGCCCTTGGCATTTCCTCAGGAGCGAATTACTTGG	3733
Qy	5603	CGTCTCATACCAAGGATAAATTAAAGGTCATTTTGCTCAAGGGAACCTCGTGAAGGAGT	5662
Db	3734	CGTCTCATACCAAGGATAAATTAAAGGTCATTTTGCTCAAGGGAACCTCGTGAAGGAGT	3793
Qy	5663	CCGCACTGAACACCAACCGGGGCGGTCCACCTCCCGCAGCGCCCAACAAAGCGAGGCC	5722
Db	3794	CCGCACTGAACACCAACCGGGGCGGTCCACCTCCCGCAGCGCCCAACAAAGCGAGGCC	3853
Qy	5723	CACCCAGTACAGGAGCACATCACCAAGCGCTGGCCTCCAGCCAGCGCCGCCCGAAG	5782
Db	3854	CACCCAGTACAGGAGCACATCACCAAGCGCTGGCCTCCAGCCAGCGCCGCCCGAAG	3913
Qy	5783	GCCCAGCCACCCGCGAGAGCAAGCACACCCACCCCTACTCCGCGAGGGCGGACCGAGC	5842
Db	3914	GCCCAGCCACCCGCGAGAGCAAGCACACCCACCCCTACTCCGCGAGGGCGGACCGAGC	3973
Qy	5843	TGCGCAGGGAACAATCTCTTGGCGCGCCCTTGAGCGAGAGAAGTCCCCCGGCGCGATGC	5902
Db	3974	TGCGCAGGGAACAATCTCTTGGCGCGCCCTTGAGCGAGAGAAGTCCCCCGGCGCGATGC	4033
Qy	5903	TCAGCAGCGGAGAGACGGTCCCCCGGAGGCTGTTTGAAGACAGCAGCAGAGGGGCCGC	5962
Db	4034	TCAGCAGCGGAGAGACGGTCCCCCGGAGGCTGTTTGAAGACAGCAGCAGAGGGGCCGC	4093
Qy	5963	TGCTCGGGGAGCGGTGAGGACCCCGCTGTCCAGAGTGAACAAGGGAAGAGGCGCAGATG	6022
Db	4094	TGCTCGGGGAGCGGTGAGGACCCCGCTGTCCAGAGTGAACAAGTCTGTGGACCAAGTCTT	4153
Qy	6023	CTCTCAAGTTTTCACGGTTAACTGTCTACCTATTAT	6060
Db	4154	CAGTATTAATCTCAGCCAGAGAAAAACCAACTCTCTCATCT	4191

RESULT 14	
ABV30132/c	
ID	ABV30132 standard; cDNA; 2896 BP.
XX	
XX	
AC	ABV30132;
XX	
XX	
DT	16-SEP-2002 (first entry)
XX	
XX	
DE	Human prostate expression marker cDNA 30123.
XX	
XX	
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW	pharmacogenomic marker; gene; ss.
XX	
OS	Homo sapiens.
OS	
XX	
FN	WO200160860-A2.
XX	
XX	
PD	23-AUG-2001.

XX		20-FEB-2001; 2001WO-US0005:71.	
PF			
XX			
PR		17-FEB-2000; 2000US-O183319P.	
PR		16-MAR-2000; 2000US-O189862P.	
PR		25-MAY-2000; 2000US-O207454P.	
PR		09-JUN-2000; 2000US-O211314P.	
PR		18-JUL-2000; 2000US-O219007P.	
PR		13-DEC-2000; 2000US-O255281P.	
XX		(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
PA		Schlegel R, Endege WO, Monahan JE;	
XX			
PI		WPI; 2001-662795/76.	
XX			
DR			
XX			
PT		Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.	
PT			
PT			
XX			
PS		Claim 1; Page 6521-6522; 11750pp; English.	
XX			
CC		The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (II) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker	
XX			
SQ		Sequence 2896 BP; 525 A; 839 C; 647 G; 885 T; 0 U; 0 Other;	
Query March 39.2%; Score 2575.2; DB 5; Length 2896;			
Best Local Similarity 94.5%; Pred. No. 0;			
Matches 2776; Conservative 0; Mismatches 3; Indels 159; Gaps 4;			
QY	762	AAACAAGATGGTGAATGCCAATCTCCGATTGGGACCCCGAGTTACATGCCTCCTGAAGT	821
DB	2896	AACAAGATGGTGAATGCCAATCTCCGATTGGGACCCCGAGTTACATGCCTCCTGAAGT	2837
QY	822	GCTGACTGTGATGAACGGGGATGAAAAGGCACCCTAGCGCTGGACTGTGACTGCTGGTC	881
DB	2836	GCTGACTGTGATGAACGGGGATGAAAAGGCACCCTAGCGCTGGACTGTGACTGCTGGTC	2777
QY	882	AGTGGCGTGATTGCCCTATGAGATGATTTATGGAGATGCCCTTCGACAGGGAACCTC	941
DB	2776	AGTGGCGTGATTGCCCTATGAGATGATTTATGGAGATGCCCTTCGACAGGGAACCTC	2717
QY	942	TGCAGAACCTTCAATAACATTATGAATTTCCAGCGTTTTTGAATTTCCAGATGACCC	1001
DB	2716	TGCAGAACCTTCAATAACATTATGAATTTCCAGCGTTTTTGAATTTCCAGATGACCC	2657
QY	1002	CAAAGTGAGCAGTGACATTTCTTGATCTGATTCAAAGCTTGTTGTGGCCGCCAGAAAGAG	1061
DB	2856	CAAAGTGAGCAGTGACATTTCTTGATCTGATTCAAAGCTTGTTGTGGCCGCCAGAAAGAG	2597
QY	1062	ACTGAAAGTTGAAGGCTTTGCTGCCATTCCTTCTCTAAAAATTGACTGGAAACAACAT	1121
DB	2596	ACTGAAAGTTGAAGGCTTTGCTGCCATTCCTTCTCTAAAAATTGACTGGAAACAACAT	2537
QY	1122	TCGTAACTCTCCTCCCCCTTCGTTCCACCCCTCAAGTCTGACGATCACACCTCCAATTT	1181
DB	2536	TCGTAACTCTCCTCCCCCTTCGTTCCACCCCTCAAGTCTGACGATCACACCTCCAATTT	2477
QY	1182	TGATGAACCGAGAAGAAATTCGTGGGTTTCATCTCTCCGTGCAGCTGAGCCCCTCAGG	1241
DB	2476	TGATGAACCGAGAAGAAATTCGTGGGTTTCATCTCTCCGTGCAGCTGAGCCCCTCAGG	2417

QY 1242 CTTCTCGGTGAGAACTGCGGTTGTGGGGTTTTCGTACAGCAAGCACTGGGATTCT 1301
Db 2416 CTTCTCGGTGAGAACTGCGGTTGTGGGGTTTTCGTACAGCAAGCACTGGGATTCT 2357
QY 1302 TGGTAGATCTGA----- 1313
Db 2356 TGGTAGATCTGAGACTTTTGGGGATGGATGAAATACGTAAAAACACTTGGTCTAGTACC 2297
QY 1314 -----GTCTGTGTCTCGGTCTGACTCCCTGTCGAAGACTAGC 1353
Db 2296 TGGCACATGGAAAGCCCTTGGTGTGTCTCGGTCTGACTCCCTGTCGAAGACTAGC 2237
QY 1354 TCCATGGAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGACTCTCAGGCAAGTGT 1413
Db 2236 TCCATGGAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGACTCTCAGGCAAGTGT 2177
QY 1414 CACAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTACAGGTGGAGGCTGTG 1473
Db 2176 CACAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTACAGGTGGAGGCTGTG 2117
QY 1474 CTTAGTCAGAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGATCCCTCTCGGAGCAG 1533
Db 2116 CTTAGTCAGAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGATCCCTCTCGGAGCAG 2057
QY 1534 GACCTTGCTACCTACATCACAGAAATGACGAGTGTAAAGCGAAATTTGGAGCAAGCAGG 1593
Db 2056 GACCTTGCTACCTACATCACAGAAATGACGAGTGTAAAGCGAAATTTGGAGCAAGCAGG 1997
QY 1594 ATGAGGTGTCAGAGAGGTGACAAAGCACTGAGCTCTCCATGATATCAGAGAGCAG 1653
Db 1996 ATGAGGTGTCAGAGAGGTGACAAAGCACTGAGCTCTCCATGATATCAGAGAGCAG 1937
QY 1654 AGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTACAGGCTCAAGTGGAAAGAAATGAGG 1713
Db 1936 AGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTACAGGCTCAAGTGGAAAGAAATGAGG 1877
QY 1714 TTGATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGAGCGAGTATCTCTAC 1773
Db 1876 TTGATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGAGCGAGTATCTCTAC 1817
QY 1774 GAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATCAAGCGGAAAGCGACAGAA 1833
Db 1816 GAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATCAAGCGGAAAGCGACAGAA 1757
QY 1834 TGTGAGCAATAAATCTGTTGAAGGTAAAGATCAAGGAAAGCTGAAGTGGAGAAATATGG 1893
Db 1756 TGTGAGCAATAAATCTGTTGAAGGTAAAGATCAAGGAAAGCTGAAGTGGAGAAATATGG 1697
QY 1894 AAATCTGAGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAGCTCCAGAGAAATCTG 1953
Db 1696 AAATCTGAGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAGCTCCAGAGAAATCTG 1637
QY 1954 GAGAAGCT-----GCA 1965
Db 1636 GAGAAGCTTAAAGCCAGCACGGAGGCCACCGAGCTGCTGCAGAAATATCCGCCAGGCA 1577
QY 1966 AAGGAGCGAGCCAGAGGAGCTGGAGAGAGCTGCAGAACGAGAGATCTTCTGAAGGC 2025
Db 1576 AAGGAGCGAGCCAGAGGAGCTGGAGAGAGCTGCAGAACGAGAGATCTTCTGAAGGC 1517
QY 2026 ATCAGAAAGAGCTGGTGGAAAGCTGAGGAAACCGCCCATTTCTTGAGAAACAGGTAAAG 2085
Db 1516 ATCAGAAAGAGCTGGTGGAAAGCTGAGGAAACCGCCCATTTCTTGAGAAACAGGTAAAG 1457
QY 2086 AGACTAGAGACCATGAGCGTAGAGAAACAGACTGAAGAGTACATCCAGCAAAATCC 2145
Db 1456 AGACTAGAGACCATGAGCGTAGAGAAACAGACTGAAGAGTACATCCAGCAAAATCC 1397
QY 2146 CAACAGATCCAGAGATGGCTGATAAAATTTCTGGAGCTGAAAGAGAAACATCGGAGGCC 2205
Db 1396 CAACAGATCCAGAGATGGCTGATAAAATTTCTGGAGCTGAAAGAGAAACATCGGAGGCC 1337
QY 2206 CAAGTCTCAGCCAGCACCTAGAGTGTGACCTGTAAACAGAAAGAGCAGCACTATGA-GGA 2264

Db 1336 CAAGTCTCAGCCAGCACCTAGAAAGTGCACTGAAACAGAAAGAGCAGCACTATGAGGGA 1277
QY 2265 AAAGATTAAAGTGTGGGCAATTCAGATAAAGAAAGACTGGCTGACAAAGGAGACACTGGA 2324
Db 1276 AAAGATTAAAGTGTGGGCAATTCAGATAAAGAAAGACTGGCTGACAAAGGAGACACTGGA 1217
QY 2325 GAAATGATGTCAGAGACACAGAGGAGGCGCCATGAGAGGCGAAATTTCTCAGCGAACA 2384
Db 1216 GAAATGATGTCAGAGACACAGAGGAGGCGCCATGAGAGGCGAAATTTCTCAGCGAACA 1157
QY 2385 GAAGCGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAACAGAGGATTGTGGA 2444
Db 1156 GAAGCGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAACAGAGGATTGTGGA 1097
QY 2445 ACTGTCTGAAGCCCAATAAATTCAGCAAAATAGCAGTCTTTTACCCAAAGGAACATGAA 2504
Db 1096 ACTGTCTGAAGCCCAATAAATTCAGCAAAATAGCAGTCTTTTACCCAAAGGAACATGAA 1037
QY 2505 GGCCCAAGAGAGATGATTTCTGAATCAGCAACAGAAATTTTACCTGGAGACACAGGC 2564
Db 1036 GGCCCAAGAGAGATGATTTCTGAATCAGCAACAGAAATTTTACCTGGAGACACAGGC 977
QY 2565 TGGAGATTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAGAAAGATCAAGCAACA 2624
Db 976 TGGAGATTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAGAAAGATCAAGCAACA 917
QY 2625 AGACACAGTCACAAGAAATCGGCTGTGGAACCTGGAGACAGATTTGGGGAGGTCACTCT 2684
Db 916 AGACACAGTCACAAGAAATCGGCTGTGGAACCTGGAGACAGATTTGGGGAGGTCACTCT 857
QY 2685 AGACACAGTCACAAGAAATCGGCTGTGGAACCTGGAGACAGTACAGAGTCTACCTCT 2744
Db 856 AGACACAGTCACAAGAAATCGGCTGTGGAACCTGGAGACAGTACAGAGTCTACCTCT 797
QY 2745 GCAGGAGCGGAGTCAGATTGACAGCCCTGCAAGCTGCAAGGCGCCCTGGAGAGGCA 2804
Db 796 GCAGGAGCGGAGTCAGATTGACAGCCCTGCAAGCTGCAAGGCGCCCTGGAGAGGCA 737
QY 2805 GCTTCGCGAGCGAGACAGAGCTGGAGAGACACACAGCAAGAGCTGAAGAGGAGATCCA 2864
Db 736 GCTTCGCGAGCGAGACAGAGCTGGAGAGACACACAGCAAGAGCGGAGAGATCCA 677
QY 2865 GGCACTCAGCGCACATAGAGATGAAATCAGCGCAAAATTTGATGCTCTTCTGTAACAGCTG 2924
Db 676 GGCACTCAGCGCACATAGAGATGAAATCAGCGCAAAATTTGATGCTCTTCTGTAACAGCTG 617
QY 2925 TACTGTAAATCAGACCTCGAGGAGCAGCTAAACCCAGCTCACCGAGACACCGCTGAAT 2984
Db 616 TACTGTAAATCAGACCTCGAGGAGCAGCTAAACCCAGCTCACCGAGACACCGCTGAAT 557
QY 2985 CAAACAACAAATCTTACTTGTCCAAACACTCGATGAGGCTTCTGGCCGCAACGACGA 3044
Db 556 CAAACAACAAATCTTACTTGTCCAAACACTCGATGAGGCTTCTGGCCGCAACGACGA 497
QY 3045 GATTGTAACTCGAAGTGAAGTGGACCATCTCCGCGGGAGATCAACGAAACGAGAGAT 3104
Db 496 GATTGTAACTCGAAGTGAAGTGGACCATCTCCGCGGGAGATCAACGAAACGAGAGAT 437
QY 3105 GCAGCTTACCAGCCAGAAACGATGGAGGCTCTGAAAGCACCGTGCACCAATGCTGGA 3164
Db 436 GCAGCTTACCAGCCAGAAACGATGGAGGCTCTGAAAGCACCGTGCACCAATGCTGGA 377
QY 3165 GGAACAGGTATGATTTGGAGGCCCTTAAACGATGAGTCTGTAGAAAAAGAGCGGAGTG 3224
Db 376 GGAACAGGTATGATTTGGAGGCCCTTAAACGATGAGTCTGTAGAAAAAGAGCGGAGTG 317
QY 3225 GGAGGCTCGAGGAGCGCTCTGGGTGATGAGAAATCCAGTGTGAGTGTGCGGTTCGAGA 3284
Db 316 GGAGGCTCGAGGAGCGCTCTGGGTGATGAGAAATCCAGTGTGAGTGTGCGGTTCGAGA 257
QY 3285 GCTGACAGAAATGCTGACACCGAGAAACAGAGCAGGCGGAGGCGGATCAGCGGATCAC 3344

Db 256 GCTGCAGAGGATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCGCGATCAC 197
QY 3345 CQAGTCTGCCAGGTGGTGGAGCTGGAGTGAAGAGCAGACAGGCTGAGATTCTCGCTCT 3404
Db 196 CQAGTCTGCCAGGTGGTGGAGCTGGAGTGAAGAGCAGACAGGCTGAGATTCTCGCTCT 137
QY 3405 GCAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCTGACAGCTCAATGA 3464
Db 136 GCAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCTGACAGCTCAATGA 77
QY 3465 CQTGAGAGAGCATGCTATGCTTGAATGATGCCCGAGCTTACAGCAGAGCTGGA 3524
Db 76 CTTGAGA-----AGAGCTGGA 59
QY 3525 GACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAATTAACAGCAGCAG 3582
Db 58 GACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAATTAACAGCAGCAG 1
RESULT 15
ID ABA08361
AC ABA08361; standard; cDNA; 3131 BP.
XX ABA08361;
XX
XX 11-JAN-2002 (first entry)
XX Human RHO/RAC effector homologue-encoding cDNA, SEQ ID NO:137.
DE
DE Human; cytokine; cell proliferation; tissue growth; differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.
XX
XX Homo sapiens.
XX
XX WO200157188-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US003800.
XX
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX
XX (HYSB-) HYSB INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-457740/49.
XX P-PSDB; ABB11117.
XX
XX Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
XX Claim 1; Page 387-388; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which

CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention
XX
XX Sequence 3131 BP; 861 A; 835 C; 806 G; 629 T; 0 U; 0 Other;
SQ
Query Match 39.0%; Score 2564.8; DB 4; Length 3131;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2781; Conservative 0; Mismatches 2; Indels 195; Gaps 2;
QY 2875 GCACATAGATGTAATCCAGCGCAATTTGATGCTCTTCTGTAACAGCTGTACTGTAACT 2934
Db 1 GCACATAGATGTAATCCAGCGCAATTTGATGCTCTTCTGTAACAGCTGTACTGTAACT 60
QY 2935 ACAGACCTGGAGGAGCAGCTAAACACGCTGACCGAGGACAAACGCTGAACCAACCAA 2994
Db 61 ACAGACCTGGAGGAGCAGCTAAACACGCTGACCGAGGACAAACGCTGAACCAACCAA 120
QY 2995 AACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGCGCCCAACGACGAGATTGTACAA 3054
Db 121 AACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGCGCCCAACGACGAGATTGTACAA 180
QY 3055 CTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATCACGGAAACGAGAGATGCAGCTTACC 3114
Db 181 CTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATCACGGAAACGAGAGATGCAGCTTACC 240
QY 3115 AGCCAGAACCAACGATGGAGGCTCTGAAGACCAACGTCGCACATGCTGGAGGAAACAGGTC 3174
Db 241 AGCCAGAACCAACGATGGAGGCTCTGAAGACCAACGTCGCACATGCTGGAGGAAACAGGTC 300
QY 3175 ATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAAAAAGAGCGGCGAGTGGAGGCTCGG 3234
Db 301 ATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAAAAAGAGCGGCGAGTGGAGGCTCGG 360
QY 3235 AGGAGCGTCTCTGGGTGATGAGAAATCCAGTTTGAAGTGTCTGGGTTCGAGAGCTGCAGAGA 3294
Db 361 AGGAGCGTCTCTGGGTGATGAGAAATCCAGTTTGAAGTGTCTGGGTTCGAGAGCTGCAGAGA 420
QY 3295 ATGCTGACACCGAGAAACAGAGCAGGGCGAGGCGATCAGCGGATCACCGAGTCTCGC 3354
Db 421 ATGCTGACACCGAGAAACAGAGCAGGGCGAGGCGATCAGCGGATCACCGAGTCTCGC 480
QY 3355 CAGGTGTGGAGCTGGCAGTGAAGAGCACAAGAGCTGAGATTCTTCGCTGTCAGCAGGCT 3414
Db 481 CAGGTGTGGAGCTGGCAGTGAAGAGCACAAGAGCTGAGATTCTTCGCTGTCAGCAGGCT 540

QY 3415 CTCAAGAGCAGAGCTGAAGGCGGAGAGCTCTCTGACAAAGCTCAATGACCTGGAGAG 3474
DB 541 CTCAAGAGCAGAGCTGAAGGCGGAGAGCTCTCTGACAAAGCTCAATGACCTGGAGAG 600
QY 3475 AAGCATCTATGCTTGAATGAATGCCGAAGCTTACAGCAGAGCTGGAGACTGAACGA 3534
DB 601 AAGCATCTATGCTTGAATGAATGCCGAAGCTTACAGCAGAGCTGGAGACTGAACGA 660
QY 3535 GAGCTCAACACAGAGGCTTCTGGAAGAGCAGCAAGCAAAATTACAGCAGCAGATGGACCTGGAG 3594
DB 661 GAGCTCAACACAGAGGCTTCTGGAAGAGCAGCAAGCAAAATTACAGCAGCAGATGGACCTGGAG 720
QY 3595 AAAAATCACATTTTCGCTCTGA CTCGAAGACTGCAAGAGCTCTAGATCGGGCTGATCTA 3654
DB 721 AAAAATCACATTTTCGCTCTGA CTCGAAGACTGCAAGAGCTCTAGATCGGGCTGATCTA 780
QY 3655 CTGAAGACAGAAAGTGAAGTCTGAGTATCAGTGGAAACATTCAGGTTCTCTATTCT 3714
DB 781 CTGAAGACAGAAAGTGAAGTCTGAGTATCAGTGGAAACATTCAGGTTCTCTATTCT 840
QY 3715 CATGAAGAGTGAATGGAAGCACTATTCTCAACAAACCAAACTCAATTGTTTCTG 3774
DB 841 CATGAAGAGTGAATGGAAGCACTATTCTCAACAAACCAAACTCAATTGTTTCTG 900
QY 3775 CAAGCCAAATGACCAACCTGCTAAAGAAAGAGGTTTATTAGTCGACGGAAGAG 3834
DB 901 CAAGCCAAATGACCAACCTGCTAAAGAAAGAGGTTTATTAGTCGACGGAAGAG 934
QY 3835 GAOCCTGCTTTACCCACACAGAGTCTCTCTGCACTACAATGAGTGAAGTGGCCCTGGAG 3894
DB 935 -----AGGTTCTCTGCACTACAATGAGTGAAGTGGCCCTGGAG 975
QY 3895 AAGGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCCCTTCAGAAAGACCGCGATCGAGCTC 3954
DB 976 AAGGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCCCTTCAGAAAGACCGCGATCGAGCTC 1035
QY 3955 CGGTCCGCGGAGAGAGTCCCAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4014
DB 1036 CGGTCCGCGGAGAGAGTCCCAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1095
QY 4015 CCAGCAACCGCGAGGAGCAGATGCGCATGTGCGCATGTGCGCATGTGCGCATGTGCGCATGTGCG 4074
DB 1096 CCAGCAACCGCGAGGAGCAGATGCGCATGTGCGCATGTGCGCATGTGCGCATGTGCGCATGTGCG 1155
QY 4075 CCCAGTGCCATGAGCTGTGCGCCCGCCATCCAGCCGAGAAAGAGTCTCAACTCA 4134
DB 1156 CCCAGTGCCATGAGCTGTGCGCCCGCCATCCAGCCGAGAAAGAGTCTCAACTCA 1215
QY 4135 GAGGAATTTAGTCCGGCTCTTAAGGAACGATGCACCAATATTCTCCACGATTTCAAC 4194
DB 1216 GAGGAATTTAGTCCGGCTCTTAAGGAACGATGCACCAATATTCTCCACGATTTCAAC 1275
QY 4195 GTAGGACTGAATGCGGACCAAGT 4254
DB 1276 GTAGGACTGAATGCGGACCAAGT 1335
QY 4255 CGCAGGCAATCCAAATGTCTGCAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4314
DB 1336 CGCAGGCAATCCAAATGTCTGCAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1395
QY 4315 TTGCGAGCACTCGCGCTTGT 4374
DB 1396 TTGCGAGCACTCGCGCTTGT 1455
QY 4375 CGTGACAAATGAATCCCGAGTCTCCAGACCAAGGAGCCAGAGAGCTTTGACCTGT 4434
DB 1456 CGTGACAAATGAATCTCCAGAGTCTCCAGACCAAGGAGCCAGAGAGCTTTGACCTGT 1515
QY 4435 GAAGGTGTGATGAGTGTGCGGAGGAAATACAAACGAGGACAGCAAGGCTGGAGAGGAG 4494
DB 1516 GAAGGTGTGATGAGTGTGCGGAGGAAATACAAACGAGGAGCAAGGCTGGAGAGGAG 1575

QY 4495 TACATTGCTCTGGAGGATCAAAAGTCTCTCAATTATGACAAATGAAGCCAGAGAGCTGGA 4554
DB 1576 TACATTGCTCTGGAGGATCAAAAGTCTCTCAATTATGACAAATGAAGCCAGAGAGCTGGA 1635
QY 4555 CAGAGCCCGTGGAGAAATTTGAGCTGTGCTTCCCGACGGGAGTGTATCTATTTCATGGT 4614
DB 1636 CAGAGCCCGTGGAGAAATTTGAGCTGTGCTTCCCGACGGGAGTGTATCTATTTCATGGT 1695
QY 4615 GCGGTGTGTGCTTCCGAACTCGCAAAATACAGCCAAAGCA ----- 4653
DB 1696 GCGGTGTGTGCTTCCGAACTCGCAAAATACAGCCAAAGCAAGTGTCTCCATACATACTGAAG 1755
QY 4654 ----- 4653
DB 1756 ATGGAATCTACCCGCGACACACACCTGCTGGCCCGGAGAAACCTCTACTTTGCTAGCTCCC 1815
QY 4654 ----- 4653
DB 1816 AGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTAGAATCAGTTGTCTCGCAGGTGGAGA 1875
QY 4654 -----GAAAGAGCAGAAAGCTGATGCTTAAACTGCTTGGAAACTCCCTGCTGAAACTG 4704
DB 1876 GTTTCTAGGGAAAAAGCAGAAAGCTGATGCTTAAACTGCTTGGAAACTCCCTGCTGAAACTG 1935
QY 4705 GAAGGTGATGACCGCTTAGACATGAACTGCAAGCTGCTCCCTTCAAGTACCAAGGTGGTGTG 4764
DB 1936 GAAGGTGATGACCGCTTAGACATGAACTGCAAGCTGCTCCCTTCAAGTACCAAGGTGGTGTG 1995
QY 4765 GTGGCCACCGAGGAAGGGCTCTACGCCCTGAAATGTCTTGAAGAACTCCCTAAACCACTGTC 4824
DB 1996 GTGGCCACCGAGGAAGGGCTCTACGCCCTGAAATGTCTTGAAGAACTCCCTAAACCACTGTC 2055
QY 4825 CCAGGAATGGAGAGCTTCTCCAAATTTATATTAAGGACCTGGAGAGCTACTCATG 4884
DB 2056 CCAGGAATGGAGAGCTTCTCCAAATTTATATTAAGGACCTGGAGAGCTACTCATG 2115
QY 4885 ATACAGAGAGAGCGGCACTGTGTCTTGTGAGCTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4944
DB 2116 ATACAGAGAGAGCGGCACTGTGTCTTGTGAGCTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2175
QY 4945 GCCCAGTCCCACTGCTGCTGCGCCGAGCAATCTACCCCAACATTTTGAAGCTGTCAAG 5004
DB 2176 GCCCAGTCCCACTGCTGCTGCGCCGAGCAATCTACCCCAACATTTTGAAGCTGTCAAG 2235
QY 5005 GGCTGCCACTTTGTTGGGCGAGCAGATGAGAAAGGGCTCTGCACTCTGCGAGCCATG 5064
DB 2236 GGCTGCCACTTTGTTGGGCGAGCAGATGAGAAAGGGCTCTGCACTCTGCGAGCCATG 2295
QY 5065 CCCAGCAAGTGTGCTATTCTCCCTACACGAAACCTCAGCAAAATCTGCAATCCGGAAG 5124
DB 2296 CCCAGCAAGTGTGCTATTCTCCCTACACGAAACCTCAGCAAAATCTGCAATCCGGAAG 2355
QY 5125 GAGATGAGACCTCAGAGCCCTGCGCTGATCCACTTCAACCAATTTACAGTATCTCTCAT 5184
DB 2356 GAGATGAGACCTCAGAGCCCTGCGCTGATCCACTTCAACCAATTTACAGTATCTCTCAT 2415
QY 5185 GGAACCAATTAATTTCTACGAAATCGCATGAAGCAGTACACGCTCGAGGAATTTCTGGAT 5244
DB 2416 GGAACCAATTAATTTCTACGAAATCGCATGAAGCAGTACACGCTCGAGGAATTTCTGGAT 2475
QY 5245 AAGATGACCATTTCTTGGCACTGCTGTTGTTGGCCCTCTTCCACAGCTTCCCTGTC 5304
DB 2476 AAGATGACCATTTCTTGGCACTGCTGTTGTTGGCCCTCTTCCACAGCTTCCCTGTC 2535
QY 5305 TCAATCTGTGAGTGAACAGCGGCGAGAGAGTACTTGTGTGTTTCCACGAA 5364
DB 2536 TCAATCTGTGAGTGAACAGCGGCGAGAGAGTACTTGTGTGTTTCCACGAA 2595
QY 5365 TTTGGAGTGTGTTGATTTCTTACGGAAGAGTGTAGCGGACAGCATCTCAAGTGGAGT 5424
DB 2596 TTTGGAGTGTGTTGATTTCTTACGGAAGAGTGTAGCGGACAGCATCTCAAGTGGAGT 2655
QY 5425 CGCTTACCTTTGGCTTTGCTTACAGAAACCTCTATCTGTTGTGACCCACTTCAACTCA 5484

Db	2656	CGCTTACCTTTGGCCTTTGGCCTTACAGAGAACCCCTATCTGTTTGTGACCCACTTCAACTCA	2715
QY	5485	CTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCAGGAGCCCTGCCGAGCGTACCTG	5544
Db	2716	CTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCAGGAGCCCTGCCGAGCGTACCTG	2775
QY	5545	GACATCCCGAACCCGCGCTACCTGGGCCCTGCCATTTCCTCAGGAGCGATTACTTGGCG	5604
Db	2776	GACATCCCGAACCCGCGCTACCTGGGCCCTGCCATTTCCTCAGGAGCGATTACTTGGCG	2835
QY	5605	TCCTCATACCGAGATAAATTAGGGTCAATTTCCTGCAAGGGNAACCTCGTGAAGGAGTCC	5664
Db	2836	TCCTCATACCGAGATAAATTAGGGTCAATTTCCTGCAAGGGNAACCTCGTGAAGGAGTCC	2895
QY	5665	GGCACTGAACACACCGGGGCCCGTCCACCTCCCGCAG	5702
Db	2896	GGCACTGAACACACCGGGGCCCGTCCACCTCCCGCAG	2933

Search completed: July 3, 2004, 20:57:55
Job time : 2262 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 11:30:09 ; Search time 15353 Seconds
(without alignments)
17387.465 Million cell updates/sec

Title: US-10-017-216-3
Perfect score: 6159
Sequence: 1 atgttgagtcacaaatagg.....ttctgagaacagattattgc 6159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues 6940544
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.nam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6159	100.0	6159	6	AX429514 Sequence
2	6159	100.0	6574	6	AX429512 Sequence
3	5668	92.0	8603	6	AX671108 Sequence
4	5666.6	92.0	6156	6	AX671112 Sequence
5	5666.4	92.0	6298	6	AX504254 Sequence
6	5666.2	92.0	6165	6	AX671105 Sequence
7	5661.4	91.9	6165	6	AX574425 Sequence
8	5650.4	91.7	6159	6	AX166510 Sequence
9	5631.1	91.4	6189	6	AX503780 Sequence
10	5629.8	91.4	6201	6	AX503778 Sequence
11	5534.8	89.9	8576	9	AX257469 Homo sapi
12	5358	87.0	5877	6	AX574427 Sequence
13	4627	75.1	6954	10	AF086824 Mus muscu
14	3602.6	58.5	5952	10	AF039218 Rattus no
15	3593.8	58.4	4967	10	AF070066 Mus muscu
16	3546.4	57.6	5019	10	MMU39904 U39904 Mus musculu
17	2436.8	39.6	5261	6	AX671110 Sequence
18	2436.8	39.6	5261	9	AB023166 Homo sapi
19	2187.4	35.5	3495	10	BC051185 Mus muscu
20	2187.4	35.5	3506	10	BC023775 Mus muscu
21	1395	22.6	1485	6	AX671037 Sequence
22	1395	22.6	1765	6	AX671044 Sequence
23	1393.4	22.6	2066	6	AX642956 Sequence
24	1391.8	22.6	1515	6	AR253937 Sequence
25	1082.4	17.6	2380	6	AX671042 Sequence
26	1082.4	17.6	2380	10	AF086823 Mus muscu
27	1074.4	17.4	1386	10	AF070085 Rattus no
28	750	12.2	1133	6	AR411948 Sequence
29	648	10.5	1058	6	AX053315 Sequence
30	592	9.6	817	6	AX053416 Sequence
31	434	7.0	2284	9	AK123136 Homo sapi
32	350.4	5.7	354	6	BD059649 Secretd
33	311.4	5.1	136436	9	AC002563 Human PAC
34	311.4	5.1	147862	9	AC079317 Homo sapi
35	311.4	5.1	190514	2	AC026363 Homo sapi
36	311.4	5.1	208686	9	AC079262 Homo sapi
37	258	4.2	258	6	AR269950 Sequence
38	257	4.2	32658	2	AC135099 Rattus no
39	257	4.2	234312	2	AC133614 Rattus no
40	257	4.2	248557	2	AC097432 Rattus no
41	256.2	4.2	207404	2	AC115795 Mus muscu
42	250	4.1	254	6	BD071469 Secretd
43	244	4.0	3835	6	AX817259 Sequence
44	244	4.0	3985	6	AX817261 Sequence
45	243.6	4.0	1530	6	AX504235 Sequence

ALIGNMENTS

RESULT 1
AX429514
LOCUS AX429514 6159 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 3 from Patent WO0234896.
ACCESSION AX429514
VERSION AX429514.1 GI:21540792
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Kapeller-Libermann,R.
TITLE 13245, a novel human myotonic dystrophy type protein kinase and
uses therefor

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0234896-A 3 02-MAY-2002;
MILLENNIUM PHARM INC (US)

FEATURES
source
Location/Qualifiers
1..6159
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 6159; DB 6; Length 6159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAATATATGAGCGCGGAATCCTTTTGGATGCTGGTCTGCTGAACCCATT 60
DB 1 ATGTTGAAGTTCAATATATGAGCGCGGAATCCTTTTGGATGCTGGTCTGCTGAACCCATT 60

QY 61 GCAGCGGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA 120
DB 61 GCAGCGGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA 120

QY 121 CAGCAGATGCTCCTCTTCCCGAGAGGGAATATAGATGCCCTCTTTGTTCTCTTTGAA 180
DB 121 CAGCAGATGCTCCTCTTCCCGAGAGGGAATATAGATGCCCTCTTTGTTCTCTTTGAA 180

QY 181 GAATGAGTCAGCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240
DB 181 GAATGAGTCAGCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240

QY 241 TCCGACACCATAGCTCAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTCAGA 300
DB 241 TCCGACACCATAGCTCAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTCAGA 300

QY 301 AGTCTTGTAGTTGTGGTCACTTTGTGTAAGTGCAGTGTGAAGAGAAAGCAACCCGGG 360
DB 301 AGTCTTGTAGTTGTGGTCACTTTGTGTAAGTGCAGTGTGAAGAGAAAGCAACCCGGG 360

QY 361 GACATCTATGCTATGAAGTGAAGAGAGGCTTTATGCGCCAGGAGAGGTTCA 420
DB 361 GACATCTATGCTATGAAGTGAAGAGAGGCTTTATGCGCCAGGAGAGGTTTCA 420

QY 421 TTTTGTGAGGAAGCGGGAACATATATCTCGAAGCACAAAGCCGCTGGATCCCCAAATTA 480
DB 421 TTTTGTGAGGAAGCGGGAACATATATCTCGAAGCACAAAGCCGCTGGATCCCCAAATTA 480

QY 481 CAGTATGCCCTTTAGGACAAAATCACTTTATCTGATGGAGGAATATCAGCTCGAGGG 540
DB 481 CAGTATGCCCTTTAGGACAAAATCACTTTATCTGATGGAGGAATATCAGCTCGAGGG 540

QY 541 GACTTGTGTCATTTTGAATAGATGAGGACAGTATAGATGAACCACTGATACAGTTT 600
DB 541 GACTTGTGTCATTTTGAATAGATGAGGACAGTATAGATGAACCACTGATACAGTTT 600

QY 601 TACTAGCTGAGCTGATTTTGGCTTCAAGCTTATCTGATGGGATACGCTGATCGA 660
DB 601 TACTAGCTGAGCTGATTTTGGCTTCAAGCTTATCTGATGGGATACGCTGATCGA 660

QY 661 GACATCAAGCTTCAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGGATTT 720
DB 661 GACATCAAGCTTCAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGGATTT 720

QY 721 GGATCTGCCGCGAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATGGGACC 780
DB 721 GGATCTGCCGCGAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATGGGACC 780

QY 781 CCAGATTAATGCTCCTGAGTCTGACTGTGATGAACCGGGATGAAAAGGCACTTAC 840
DB 781 CCAGATTAATGCTCCTGAGTCTGACTGTGATGAACCGGGATGAAAAGGCACTTAC 840

QY 841 GGCTGGACTGTGACTGGTGTGAGTGGGCGTGAATTCGCTATGAGATGATTTATGGGAGA 900
DB 841 GGCTGGACTGTGACTGGTGTGAGTGGGCGTGAATTCGCTATGAGATGATTTATGGGAGA 900

QY 901 TCCCCCTTCGAGAGGGAACCTCTGCGAGAACCTTCAATAACATTATGAATTTCCAGCGG 960
DB 901 TCCCCCTTCGAGAGGGAACCTCTGCGAGAACCTTCAATAACATTATGAATTTCCAGCGG 960

QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATCAAGC 1020
DB 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATCAAGC 1020

QY 1021 TTGTTGCGGCCAGAAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTCTTTCTTC 1080
DB 1021 TTGTTGCGGCCAGAAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTCTTTCTTC 1080

QY 1081 TCTAAAATTGATCGGAACAACATTCTGTAATCTCTCTCCCTCCCTTCGTTCCACCTCAAG 1140
DB 1081 TCTAAAATTGATCGGAACAACATTCTGTAATCTCTCTCCCTCCCTTCGTTCCACCTCAAG 1140

QY 1141 TCTGACGATGACCTCCAAATTTTGTGAACAGAGAGAAATTCGTGGGTTTCACTCTCT 1200
DB 1141 TCTGACGATGACCTCCAAATTTTGTGAACAGAGAGAAATTCGTGGGTTTCACTCTCT 1200

QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260
DB 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260

QY 1261 TACAGCAGGCACTGGGATTCCTGTTAGATCTGAGTCTGTTGTCGGGTCGGACTCC 1320
DB 1261 TACAGCAGGCACTGGGATTCCTGTTAGATCTGAGTCTGTTGTTGTCGGGTCGGACTCC 1320

QY 1321 CCTGCCAAGCTAGCTCCATGGAAGAAAGAACTTCTCATAAAAGCAAGAGCTACAAGAC 1380
DB 1321 CCTGCCAAGCTAGCTCCATGGAAGAAAGAACTTCTCATAAAAGCAAGAGCTACAAGAC 1380

QY 1381 TCTCAGCAAGTGTCAAGAATGAGCAGAAATGACCCGTTTACATCGAGAGTGTCA 1440
DB 1381 TCTCAGCAAGTGTCAAGAATGAGCAGAAATGACCCGTTTACATCGAGAGTGTCA 1440

QY 1441 GAGTGTGAGGCTGTGTTAGTCAAGAGGAGTGAAGCTCTGAGACTCAGAGA 1500
DB 1441 GAGTGTGAGGCTGTGTTAGTCAAGAGGAGTGAAGCTCTGAGACTCAGAGA 1500

QY 1501 TCCCTCTCGAGCAGGACCTTGTCTA CTTA CAT CACAGATGCAAGTGTCTTAAAGCAAGT 1560
DB 1501 TCCCTCTCGAGCAGGACCTTGTCTA CTTA CAT CACAGATGCAAGTGTCTTAAAGCAAGT 1560

QY 1561 TTGAGCAGCAGCGATGAGGTTGCCAGAGAGTGCACAAACACCTGACGCTTCTCCAT 1620
DB 1561 TTGAGCAGCAGCGATGAGGTTGCCAGAGAGTGCACAAACACCTGACGCTTCTCCAT 1620

QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGCTCAA 1680
DB 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGCTCAA 1680

QY 1681 GTGGAAGAAATGAGTTGATGATCAATCAGTCTGGAAGAGATCTTGTCTCAGCAAGA 1740
DB 1681 GTGGAAGAAATGAGTTGATGATCAATCAGTCTGGAAGAGATCTTGTCTCAGCAAGA 1740

QY 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800
DB 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800

QY 1801 CGGAAGCGCAGAAATGTGACATAAATCTGTTGAAGCTTAAGGATCAAGGAGGCTGAA 1860
DB 1801 CGGAAGCGCAGAAATGTGACATAAATCTGTTGAAGCTTAAGGATCAAGGAGGCTGAA 1860

QY 1861 GTGGGAGAAATATGGAATCTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
DB 1861 GTGGGAGAAATATGGAATCTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920

QY 1921 CTCAAAGAGAAATCTGGAGAGGCTGAAAGAGCGGAGGAGGCTGGAGAGCTG 1980
DB 1921 CTCAAAGAGAAATCTGGAGAGGCTGAAAGAGCGGAGGAGGCTGGAGAGCTG 1980

QY 1981 CAGAACCGAGGAGTCTTCTGTAAGGCTCAGAAAAGAGCTGTGGAAGCTGAGGAACGC 2040

1981 CAGAACCGAGAGGATTCCTCTGAAGCATCAGAAAGAGCTGGTGAAGCTGAGGAACGC 2040
2041 CCGCATTCCTCTGGAGAAACAGGTAAAGAGCTAGAGACCATGAGCGGTAGAGAAACAGA 2100
2041 CCGCATTCCTCTGGAGAAACAGGTAAAGAGCTAGAGACCATGAGCGGTAGAGAAACAGA 2100
2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTCGT 2160
2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTCGT 2160
2161 GAGCTCGAAGAGAAACATCGGAGGCCCCAAGTCTCAGCCCCAGCACCTAGAGAGTGACCTG 2220
2161 GAGCTCGAAGAGAAACATCGGAGGCCCCAAGTCTCAGCCCCAGCACCTAGAGAGTGACCTG 2220
2221 AAACAGAAAGACGACCATATCAGGAAAGATTAAGTGTGGACAATCAGATAAAGAA 2280
2221 AAACAGAAAGACGACCATATGAGAAAGATTAAGTGTGGACAATCAGATAAAGAA 2280
2281 GACCTGGCTGACAAGGAGACATCGGAGAACATGATGCAAGACACGAGGAGAGGCCCAT 2340
2281 GACCTGGCTGACAAGGAGACATCGGAGAACATGATGCAAGACACGAGGAGAGGCCCAT 2340
2341 GAGAGGCGAAATTCAGCGAAACAGAGGCGATGATCAATGCTATGGATCCAGATC 2400
2341 GAGAGGCGAAATTCAGCGAAACAGAGGCGATGATCAATGCTATGGATCCAGATC 2400
2401 AGATCCCTGGAACAGAGATTCTGGAACCTGTCTGAAGCCAAATAAATTTGCAACAAATAGC 2460
2401 AGATCCCTGGAACAGAGATTCTGGAACCTGTCTGAAGCCAAATAAATTTGCAACAAATAGC 2460
2461 AGTCTTTTACCCAAAGAAACATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520
2461 AGTCTTTTACCCAAAGAAACATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520
2521 CAGAAATTTACTCGAGACACAGCTGGAGATTTGGAGGCCAGAAACCGAATCTGGAG 2580
2521 CAGAAATTTACTCGAGACACAGCTGGAGATTTGGAGGCCAGAAACCGAATCTGGAG 2580
2581 GAGCAGCTGGAGAGATCAGCCACCAAGACCACTGACAGAAATCGGCTCGTGAACCTG 2640
2581 GAGCAGCTGGAGAGATCAGCCACCAAGACCACTGACAGAAATCGGCTCGTGAACCTG 2640
2641 GAGCAAGATTTGCGGGAGTCTAGTCTAGACACGAGGAGCAGAACTGGAGCTCAGGCG 2700
2641 GAGCAAGATTTGCGGGAGTCTAGTCTAGACACGAGGAGCAGAACTGGAGCTCAGGCG 2700
2701 CAGCTCAGAGACTACAGCTCTCCCTGAGGAGCGGAGTCACTGACAGCCCTGCGAG 2760
2701 CAGCTCAGAGACTACAGCTCTCCCTGAGGAGCGGAGTCACTGACAGCCCTGCGAG 2760
2761 GCTGACGCGGCGCTGAGAGCCAGCTTTCGACAGGCGGAGTCACTGACAGCCCTGCGAG 2820
2761 GCTGACGCGGCGCTGAGAGCCAGCTTTCGACAGGCGGAGTCACTGACAGCCCTGCGAG 2820
2821 ACAGCAGAGCTGAAGAGAGATCCAGGACCTCAGGACATAGAGATGAATCCAGGCG 2880
2821 ACAGCAGAGCTGAAGAGAGATCCAGGACCTCAGGACATAGAGATGAATCCAGGCG 2880
2881 AAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCAGACCTGGAGGAGCAGCTAAAC 2940
2881 AAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCAGACCTGGAGGAGCAGCTAAAC 2940
2941 CAGCTGACCGGAGCAACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTC 3000
2941 CAGCTGACCGGAGCAACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTC 3000
3001 GATGAGGCTTCTGGCGCCCAACGAGATTTGACAACTGCGAAGTGAAGTGACCATCTC 3060
3001 GATGAGGCTTCTGGCGCCCAACGAGATTTGACAACTGCGAAGTGAAGTGACCATCTC 3060
3061 CCGCGGAGATCAGCGAACGAGATGCAAGTTTACAGCCAGAACCAACGATGAGGCT 3120

3061 CCGCGGAGATCAGCGAACGAGAGATGCAAGTTTACAGCCAGAAAGCAACGATGAGGCT 3120
3121 CTGAAGACCACTGTCACCATGCTGGAGGAACAGGTCTATGATTTGGAGGCCCTTAAACGAT 3180
3121 CTGAAGACCACTGTCACCATGCTGGAGGAACAGGTCTATGATTTGGAGGCCCTTAAACGAT 3180
3181 GAGCTGTAGAAAAAGAGCGGAGCTGAGGAGCTTGGAGGAGCGCTCTCTGGGTGATGAGAAA 3240
3181 GAGCTGTAGAAAAAGAGCGGAGCTGAGGAGCGCTTGGAGGAGCGCTCTCTGGGTGATGAGAAA 3240
3241 TCCAGTTTGAAGTGTCCGGTTCCAGAGCTGCGAGAGATGCTGGACACCCGAGAAACAGAGC 3300
3241 TCCAGTTTGAAGTGTCCGGTTCCAGAGCTGCGAGAGATGCTGGACACCCGAGAAACAGAGC 3300
3301 AGGCGAGAGCCGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3360
3301 AGGCGAGAGCCGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3360
3361 GAGCACAAGGCTGAGATTTCTCGTCTGACAGAGCTCTCAAGAGCAGAAAGCTGGAAGGCC 3420
3361 GAGCACAAGGCTGAGATTTCTCGTCTGACAGAGCTCTCAAGAGCAGAAAGCTGGAAGGCC 3420
3421 GAGAGCCTCTCTGACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3480
3421 GAGAGCCTCTCTGACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3480
3481 GCCCGAAGCTTACAGCAGAAAGCTGGAGACTGAAAGAGCTCAACACAGAGCTTCTGGAA 3540
3481 GCCCGAAGCTTACAGCAGAAAGCTGGAGACTGAAAGAGCTCAACACAGAGCTTCTGGAA 3540
3541 GAGCAAGCCAAATTAAGCAGCAGATGGAGCTCGAGAGAAATCACAATTTCCCTCTGACT 3600
3541 GAGCAAGCCAAATTAAGCAGCAGATGGAGCTCGAGAGAAATCACAATTTCCCTCTGACT 3600
3601 CAAGAGCTGCAAGAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660
3601 CAAGAGCTGCAAGAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660
3661 GAGTATCAGCTGGAAACATTCAGGTTCTCTATTCATGAAAGGTGAAATGGAAGGC 3720
3661 GAGTATCAGCTGGAAACATTCAGGTTCTCTATTCATGAAAGGTGAAATGGAAGGC 3720
3721 ACTATTTCTCAACAAACCAAACTCATTGATTTCTGCAAGCCAAATGGAACCACTGCT 3780
3721 ACTATTTCTCAACAAACCAAACTCATTGATTTCTGCAAGCCAAATGGAACCACTGCT 3780
3781 AAAAGAAAAAGGTTTATTTAGTCGACGGAAGAGGACCCCTCTTTACCCACACAGGTT 3840
3781 AAAAGAAAAAGGTTTATTTAGTCGACGGAAGAGGACCCCTGCTTTACCCACACAGGTT 3840
3841 CCTCTCAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTCAGAG 3900
3841 CCTCTCAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTCAGAG 3900
3901 CTAGAGAAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCCGCCGGAGGAGGCTGCC 3960
3901 CTAGAGAAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCCGCCGGAGGAGGCTGCC 3960
3961 CACCGAAAAGCAACGAGCAACCCACCCATCCAGCCAGCCGCGGAGGAGCAGATC 4020
3961 CACCGAAAAGCAACGAGCAACCCACCCATCCAGCCAGCCGCGGAGGAGCAGATC 4020
4021 GCCATGTCGCGCATCGTGGGTGCGCAGAGCAGCCAGCCAGTCCCATGAGCCTGCTGGCC 4080
4021 GCCATGTCGCGCATCGTGGGTGCGCAGAGCAGCCAGCCAGTCCCATGAGCCTGCTGGCC 4080
4081 CCGCCATCCAGCGCGAGAAAGGATCTTCACTCCAGAGAAATTTAGTCGGCGCTTTAAG 4140
4081 CCGCCATCCAGCGCGAGAAAGGATCTTCACTCCAGAGAAATTTAGTCGGCGCTTTAAG 4140
4141 GAACGATGACCAACAATATTCCTCAGGATTCAGTGGAGTGAACATGCGAGCCACA 4200
4141 GAACGATGACCAACAATATTCCTCAGGATTCAGTGGAGTGAACATGCGAGCCACA 4200

[illegible]

Qy	5281	CCTGTCCTCAATCGTGCAGGTGAACAGCGCAGGCGCAGGAGAGTACTTGCCTGTGTTTC	5340
Db	5281	CCTGTCCTCAATCGTGCAGGTGAACAGCGCAGGCGCAGGAGAGTACTTGCCTGTGTTTC	5340
Qy	5341	CAGCAATTTCCGAGTGTTCGTGGATTCTTACGGAAGACGTAGCCGCACAGACGATCTCAAG	5400
Db	5341	CACGAATTTCCGAGTGTTCGTGGATTCTTACGGAAGACGTAGCCGCACAGACGATCTCAAG	5400
Qy	5401	TGGAGTCGCTTACCTTTTGGCCTTTTGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC	5460
Db	5401	TGGAGTCGCTTACCTTTTGGCCTTTTGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC	5460
Qy	5461	AACCTACTCGAAGTAATTGAGATCCAGGACGCTCTCAGCAGGGACCCCTGCCCGCAGCG	5520
Db	5461	AACCTACTCGAAGTAATTGAGATCCAGGACGCTCTCTCAGAGGACCCCTGCCCGCAGCG	5520
Qy	5521	TACTGACATCCGAAACCCGCGCTACCTTGGGCCCTGCCATTTCTCAGAGCGATTTAC	5580
Db	5521	TACTGACATCCGAAACCCGCGCTACCTTGGGCCCTGCCATTTCTCAGAGCGATTTAC	5580
Qy	5581	TTGGCGTCTCATACAGGATAAATTAAGGGTCAATTTGCTGCAAGGAAACCTCTCGTAG	5640
Db	5581	TTGGCGTCTCATACAGGATAAATTAAGGGTCAATTTGCTGCAAGGAAACCTCTCGTAG	5640
Qy	5641	GAGTCCGGCACTGAACACCAACCGGGGCCGTCCACCTCCCGCAGACGCCCCAAACAAGCGA	5700
Db	5641	GAGTCCGGCACTGNAACACACCGGGGCCGTCCACCTCCCGCAGACGCCCCAAACAAGCGA	5700
Qy	5701	GGCCACCCACGTAAACAGAGACATCAACAAGCGGTGGCTTCCAGCCGAGCGCCGCC	5760
Db	5701	GGCCACCCACGTAAACAGAGACATCAACAAGCGGTGGCTTCCAGCCGAGCGCCGCC	5760
Qy	5761	GAAGGCCCCAGCCACCGCGAGAGCCAAAGCACACCCACCGCTTACCGCAGGGGGCGGAC	5820
Db	5761	GAAGGCCCCAGCCACCGCGAGAGCCAAAGCACACCCACCGCTTACCGCAGGGGGCGGAC	5820
Qy	5821	GAGTCGCGAGGACAAAGTCTCTGCGCGCCCTCTGGAGGAGAGAAGTCTCCCGCGCCG	5880
Db	5821	GAGTCGCGAGGACAAAGTCTCTGCGCGCCCTCTGGAGGAGAGAAGTCTCCCGCGCCG	5880
Qy	5881	ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTTCAAGACACAGCAGCGGGC	5940
Db	5881	ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACACAGCAGCGGGC	5940
Qy	5941	CGGCTGCTCGGGAGCGGTGAGAACCCCGCTCTCCAGGTGAACAAGGAAAGAGGGCAG	6000
Db	5941	CGGCTGCTCGGGAGCGGTGAGAACCCCGCTCTCCAGGTGAACAAGGAAAGAGGGCAG	6000
Qy	6001	AGTGCCTCTCAAGTTTACGGTTACACTGTACCTATTATGACTGGATATAAAGCTG	6060
Db	6001	AGTGCCTCTCAAGTTTACGGTTTACACTGTACCTATTATGACTGGATATAAAGCTG	6060
Qy	6061	GACAACCTGCCAGCTAACTGGTCAGTCTTGAGGATCATCCAGCTGAATGGAGAAATCCGG	6120
Db	6061	GACAACTGCCAGCTAACTGGTCAGTCTTGAGGATCATCCAGCTGAATGGAGAAATCCGG	6120
Qy	6121	CAGCAGGTTGAAAAGTCTGTTCTGAGAACAGATTATTGC	6159
Db	6121	CAGCAGGTTGAAAAGTCTGTTCTGAGAACAGATTATTGC	6159

RESULT 2					
AX429512					
LOCUS				DNA	linear
DEFINITION	Sequence 1 from Patent WO0234896.		6574 bp		PAT 21-JUN-2002
ACCESSION	AX429512				
VERSION	AX429512.1		GI:21540791		
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				

AUTHORS Kapeller-Libermann,R.
TITLE 13245, a novel human myotonic dystrophy type protein kinase and
uses therefor
JOURNAL Patent: WO 0234896-A 1 02-MAY-2002;
MILLENNIUM PHARM INC (US)
FEATURES Location/Qualifiers
Source 1..6574
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 6159; DB 6; Length 6574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGTGTGCTGCTGAACCCATT 60
Db 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGTGTGCTGCTGAACCCATT 78
QY 61 GCACGCGGCTCCAGGCTGAATCTGTTCTCCGCGGGAACCACTTTATGACTCAA 120
Db 79 GCACGCGGCTCCAGGCTGAATCTGTTCTCCAGGGAACCACTTTATGACTCAA 138
QY 121 CAGCAGATGTCCTCTCTTTCCGAGAAAGGATATTAGATGCCCTCTTTGTTCTTTTGA 180
Db 139 CAGCAGATGTCCTCTCTTTCCGAGAAAGGATATTAGATGCCCTCTTTGTTCTTTGA 198
QY 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAAGCAAGTGAAGCACTTTGTCGGAAATAT 240
Db 199 GAATGAGTCAAGCTGCTCTGATGAAGATTAAGCAAGTGAAGCACTTTGTCGGAAATAT 258
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCAAGGACTTCGAAGTCAGA 300
Db 259 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCAAGGACTTCGAAGTCAGA 318
QY 301 AGTCTGTAGTTGGTGGTCACTTTGCTGAAGTGCAGGTGGTGAAGAGAGAAACACCGG 360
Db 319 AGTCTGTAGTTGGTGGTCACTTTGCTGAAGTGCAGGTGGTGAAGAGAGAAACACCGG 378
QY 361 GACATCTAGTATGAAGTGAATGAAGAGAGGCTTTATTTGGCCCGAGGACAGGTTTCA 420
Db 379 GACATCTAGTATGAAGTGAATGAAGAGAGGCTTTATTTGGCCCGAGGACAGGTTTCA 438
QY 421 TTTTGTGAGAAAGCGGAACATATTATCTCGAAGCAACAAGCCCGTGGATCCCCCAATTA 480
Db 439 TTTTGTGAGAAAGCGGAACATATTATCTCGAAGCAACAAGCCCGTGGATCCCCCAATTA 498
QY 481 CAGTATGCCCTTTCAGGAACAATAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGG 540
Db 499 CAGTATGCCCTTTCAGGAACAATAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGG 558
QY 541 GACTTGCTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATAACCTGATACAGTTT 600
Db 559 GACTTGCTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATAACCTGATACAGTTT 618
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTACACGGTTTCATCTGATGGGATACGTCATCGA 660
Db 619 TACCTAGCTGAGCTGATTTTGGCTGTTACACGGTTTCATCTGATGGGATACGTCATCGA 678
QY 661 GACATCAAGCCTGAGAACATTTCTCGTTGACCCGACAGGACACATCAAGCTGGTGAATTT 720
Db 679 GACATCAAGCCTGAGAACATTTCTCGTTGACCCGACAGGACACATCAAGCTGGTGAATTT 738
QY 721 GGATCTGCCCGGAAATGAATTAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
Db 739 GGATCTGCCCGGAAATGAATTAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 798
QY 781 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGAACCGGGATGGAAGGACCTTAC 840
Db 799 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGAACCGGGATGGAAGGACCTTAC 858
QY 841 GGCCTGGACTGTGACTGGTGGTCACTGGGCGGTGATTCCTATGAGATGATTTATGGGAGA 900

Db 859 GGCCTGGACTGTGACTGGTGGTCACTGGTGGCGGTGATTCCTATGAGATGATTTATGGAGA 918
QY 901 TCCCCCTTCGACAGAGGAACTCTTCGCCAGAACCTTCAATAACATTAATGAATTTCCACGG 960
Db 919 TCCCCCTTCGACAGAGGAACTCTTCGCCAGAACCTTCAATAACATTAATGAATTTCCACGG 978
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
Db 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1038
QY 1021 TTGTTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1080
Db 1039 TTGTTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1098
QY 1081 TCTAAATTTGACTGGAAACAACATTCGTAACCTCTCTCCCCCTTGGTTCACACCTCAAG 1140
Db 1099 TCTAAATTTGACTGGAAACAACATTCGTAACCTCTCTCCCCCTTGGTTCACACCTCAAG 1158
QY 1141 TCTGACCATGACACCTCCAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTCATCTCT 1200
Db 1159 TCTGACCATGACACCTCCAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTCATCTCT 1218
QY 1201 CGGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCGTTTGGGGTTTTTCG 1260
Db 1219 CGGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCGTTTGGGGTTTTTCG 1278
QY 1261 TACAGCAAGCACTGGGATTTCTTGTGTAGATCTGAGTCTGTTGTCGGGTCTGGAATCC 1320
Db 1279 TACAGCAAGCACTGGGATTTCTTGTGTAGATCTGAGTCTGTTGTCGGGTCTGGAATCC 1338
QY 1321 CCTGCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAGCAAAAGAGCTACAAGAC 1380
Db 1339 CCTGCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAGCAAAAGAGCTACAAGAC 1398
QY 1381 TCTCAGACAAGTGTCAAAAGATGGACAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1440
Db 1399 TCTCAGACAAGTGTCAAAAGATGGACAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1458
QY 1441 GAGGTGAGGCTGTGCTTAGTCAAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1459 GAGGTGAGGCTGTGCTTAGTCAAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1518
QY 1501 TCCCTCTCGACAGGACCTTGTCTACATCAAGAATGCAAGTGAAGTGAAGTGAAGTGAAGT 1560
Db 1519 TCCCTCTCGACAGGACCTTGTCTACATCAAGAATGCAAGTGAAGTGAAGTGAAGTGAAGT 1578
QY 1561 TTGGAGCAAGCAGGATGGAGGTGTCCTCCAGAGGATGACAAAGCACTGCAAGCTTCTCCAT 1620
Db 1579 TTGGAGCAAGCAGGATGGAGGTGTCCTCCAGAGGATGACAAAGCACTGCAAGCTTCTCCAT 1638
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1680
Db 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1698
QY 1681 GTGAAAGAAATGAGGTGATGATGAATCAGTTTGAAGAGGATCTGTTCTCAGCAAGAGA 1740
Db 1699 GTGAAAGAAATGAGGTGATGATGAATCAGTTTGAAGAGGATCTGTTCTCAGCAAGAGA 1758
QY 1741 CGGAGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1800
Db 1759 CGGAGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1818
QY 1801 CGGAAACGCAAGATCTCAGATTAACCTTTGAAGGCTTAAGGATCAAGGAAAGCTGTAA 1860
Db 1819 CGGAAACGCAAGATCTCAGATTAACCTTTGAAGGCTTAAGGATCAAGGAAAGCTGTAA 1878
QY 1861 GTGGGAATATCGAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTTCAAGAG 1920
Db 1879 GTGGGAATATCGAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTTCAAGAG 1938
QY 1921 CTCACAGAGAAACTGGAGAAAGGCTGCAAGAGGACGAGCCGAGAGGAGCTGGAGAAAGCTG 1980

Db 1939 CTCRAAGAGAACTGGAGAGGCTGCAAGAGCGAGCGAGGAGCTGGAGAGCTG 1998
Qy 1981 CAGAAACGAGAGATTTCTGTAAGGCATCAGAAAGAGCTGGTGAAGCTGAGAAAGC 2040
Db 1999 CAGAAACGAGAGATTTCTGTAAGGCATCAGAAAGAGCTGGTGAAGCTGAGAAAGC 2058
Qy 2041 CGCAATCTCTGAGAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2100
Db 2059 CGCAATCTCTGAGAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2118
Qy 2101 CTGAAGGATGACATCCAGACAAATCCACAGATCCAGCAGATGCGTCTGATAAAATCTG 2160
Db 2119 CTGAAGGATGACATCCAGACAAATCCACAGATCCAGCAGATGCGTCTGATAAAATCTG 2178
Qy 2161 GAGCTCGAGAGAAACATCGGAGGCGCAAGTCTGAGCCAGCAGCTAGAGTGCACCTG 2220
Db 2179 GAGCTCGAGAGAAACATCGGAGGCGCAAGTCTGAGCCAGCAGCTAGAGTGCACCTG 2238
Qy 2221 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAA 2280
Db 2239 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAA 2298
Qy 2281 GAGCTGGTGAAGAGAGACACTGGAGAACATGATGCAGAGACACAGAGAGAGGCCAT 2340
Db 2299 GAGCTGGTGAAGAGAGACACTGGAGAACATGATGCAGAGACACAGAGAGAGGCCAT 2358
Qy 2341 GAGAAAGGCAAAATCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAAGATC 2400
Db 2359 GAGAAAGGCAAAATCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCCAAGATC 2418
Qy 2401 AGATCCCTGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAACTTGCAGCAATAGC 2460
Db 2419 AGATCCCTGAAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAACTTGCAGCAATAGC 2478
Qy 2461 AGTCTTTTACCAAGAGAACATGAGGCGCCAGAGAGAGATGATTTCTGAATCTCAGGCAA 2520
Db 2479 AGTCTTTTACCAAGAGAACATGAGGCGCCAGAGAGAGATGATTTCTGAATCTCAGGCAA 2538
Qy 2521 CAGAAATTTTACCTGGAGACACAGGCTGGGAAATTTGGAGGCCAGAACCGAAATCTGGAG 2580
Db 2539 CAGAAATTTTACCTGGAGACACAGGCTGGGAAATTTGGAGGCCAGAACCGAAATCTGGAG 2598
Qy 2581 GAGCAGCTGGAGAGATCAGCCACAGACCAAGCCAGTGAAGAAATCGGCTGTGAATCTG 2640
Db 2599 GAGCAGCTGGAGAGATCAGCCACCAAGACCAGTGAAGAAATCGGCTGTGAATCTG 2658
Qy 2641 GAGACAAGATTTGGGAGGTGAGTCTGAGACACGAGGAGCAAACTGGAGCTCAAGGCG 2700
Db 2659 GAGACAAGATTTGGGAGGTGAGTCTGAGACACGAGGAGCAAACTGGAGCTCAAGGCG 2718
Qy 2701 CAGCTCACAGAGTACAGCTCTCCCTGCGAGGCGGAGTCAAGTGTGACGCGCTGAG 2760
Db 2719 CAGCTCACAGAGTACAGCTCTCCCTGCGAGGCGGAGTCAAGTGTGACGCGCTGAG 2778
Qy 2761 GCTGACGCGCGCCCTGGAGAGCCAGTTCGCGAGGCAAGACAGAGCTGGAAGAGACC 2820
Db 2779 GCTGACGCGCGCCCTGGAGAGCCAGTTCGCGAGGCAAGACAGAGCTGGAAGAGACC 2838
Qy 2821 ACAGCAGAGCTGAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATTCAGGCG 2880
Db 2839 ACAGCAGAGCTGAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATTCAGGCG 2898
Qy 2881 AATTTGATGCTCTTCGTAACAGCTGTAAGTAAATCAGAGACCTGGAGAGCAGCTAAAC 2940
Db 2899 AATTTGATGCTCTTCGTAACAGCTGTAAGTAAATCAGAGACCTGGAGAGCAGCTAAAC 2958
Qy 2941 CAGCTACCGAGAGACAGCTGAACCTCAACCAAACTTCTACTTGTCCAAACAACTC 3000
Db 2959 CAGCTACCGAGAGACAGCTGAACCTCAACCAAACTTCTACTTGTCCAAACAACTC 3018
Qy 3001 GATGAGCTTCTGGCCCAACGAGAGATTTGTAACCTCGAGGTGAAGTGGACCATCTC 3060
Db 3019 GATGAGGCTTCTGGCCCAACGAGAGATTTGTAACCTCGAGGTGAAGTGGACCATCTC 3078

Qy 3061 GCCTGGAGATCAGGAAACGAGAGATCGAGCTTACCGAGCAGAGCAAAAGATGGAGCT 3120
Db 3079 GCCTGGAGATCAGGAAACGAGAGATCGAGCTTACCGAGCAGAGCAAAAGATGGAGCT 3138
Qy 3121 CTGAAGACCAAGCTGCACCATCTGGAGGAAACAGGTCAATGATTTGGAGGCCCTAAACGAT 3180
Db 3139 CTGAAGACCAAGCTGCACCATCTGGAGGAAACAGGTCAATGATTTGGAGGCCCTAAACGAT 3198
Qy 3181 GAGTGTCTAGAAAAGAGCGGAGTGGAGGCGCTGGAGGCGCTCTGGTGTGTGAGAAA 3240
Db 3199 GAGTGTCTAGAAAAGAGCGGAGTGGAGGCGCTGGAGGCGCTCTGGTGTGTGAGAAA 3258
Qy 3241 TCCAGATTTGAGTGTGGGTTTCGAGAGCTGCAGAGATGCTGACACCGAGAAACAGAGC 3300
Db 3259 TCCAGATTTGAGTGTGGGTTTCGAGAGCTGCAGAGATGCTGACACCGAGAAACAGAGC 3318
Qy 3301 AGGCGAGAGCCGATCAGCGGATCAACGAGTCTGCGCAGGTGTGAGCTGGCAGTGAAG 3360
Db 3319 AGGCGAGAGCCGATCAGCGGATCAACGAGTCTGCGCAGGTGTGAGCTGGCAGTGAAG 3378
Qy 3361 GAGCACAGGCTGAGATTTCTCGCTCTGCGCAGGCTCTCAAGAGCAGAGCTGAAGGCC 3420
Db 3379 GAGCACAGGCTGAGATTTCTCGCTCTGCGCAGGCTCTCAAGAGCAGAGCTGAAGGCC 3438
Qy 3421 GAGAGCCTCTCTGACAAAGCTCAATGACCTGGAGAAGCATGCTATGCTTGAATGAAT 3480
Db 3439 GAGAGCCTCTCTGACAAAGCTCAATGACCTGGAGAAGCATGCTATGCTTGAATGAAT 3498
Qy 3481 GCCGAGAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAA 3540
Db 3499 GCCGAGAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAA 3558
Qy 3541 GAGCAAGCCAAATTAACAGCAGCAGATGGAACCTGCAGAAAAATCACATTTTCCGCTGACT 3600
Db 3559 GAGCAAGCCAAATTAACAGCAGCAGATGGAACCTGCAGAAAAATCACATTTTCCGCTGACT 3618
Qy 3601 CAAGGATCGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGATGACTTG 3660
Db 3619 CAAGGATCGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGATGACTTG 3678
Qy 3661 GAGTATCAGCTGGAAGACATTCAGGTTCTATTCCTCATGAAAGGTGAATTTGAAGGC 3720
Db 3679 GAGTATCAGCTGGAAGACATTCAGGTTCTATTCCTCATGAAAGGTGAATTTGAAGGC 3738
Qy 3721 ACTATTTCTCAACAAACCAAACTCATTTGTTCTGCAAGCCAAAATGGAACCACTGCT 3780
Db 3739 ACTATTTCTCAACAAACCAAACTCATTTGTTCTGCAAGCCAAAATGGAACCACTGCT 3798
Qy 3781 AAAAGAAAAGGTTTATTTAGTCCGAGGAAAGAGACCTCTTTTACCCACACAGGTT 3840
Db 3799 AAAAGAAAAGGTTTATTTAGTCCGAGGAAAGAGACCTCTTTTACCCACACAGGTT 3858
Qy 3841 CCTCTGAGTACAAATGAGCTGGAAGTGGCCCTGGAGAGGAGAAAGCTCCTCTGCAGAG 3900
Db 3859 CCTCTGAGTACAAATGAGCTGGAAGTGGCCCTGGAGAGGAGAAAGCTCCTCTGCAGAG 3918
Qy 3901 CTAGAGGAAGCCCTTCAAGAGACCGGATCGGCTCGGTTCCGCGGAGGAGGAGCTGCC 3960
Db 3919 CTAGAGGAAGCCCTTCAAGAGACCGGATCGGCTCGGTTCCGCGGAGGAGGAGCTGCC 3978
Qy 3961 CACCGCAAAAGCAACGAGACCAACCAACCCATCCACGCCACCCCGAGGAGCAGATC 4020
Db 3979 CACCGCAAAAGCAACGAGACCAACCAACCCATCCACGCCACCCCGAGGAGCAGATC 4038
Qy 4021 GCCATGTCCGCATCTGTGCGGTTCGCCAGAGACACAGCCAGTCCCATGAGCTCTGGCC 4080
Db 4039 GCCATGTCCGCATCTGTGCGGTTCGCCAGAGACACAGCCAGTCCCATGAGCTCTGGCC 4098
Qy 4081 CCGGCATCCAGCGCAGAAAGAGGAGTCTTCAACTCCAGAGAAATTTAGTGGCGCTCTTAAG 4140
Db 4099 CCGGCATCCAGCGCAGAAAGAGGAGTCTTCAACTCCAGAGAAATTTAGTGGCGCTCTTAAG 4158

QY 4141 GAACGATGACCAACAATATTCCTCACCGATTCAACGTAAGTGAACATCGAGCCACA 4200
Db 4159 GAACGATGACCAACAATATTCCTCACCGATTCAACGTAAGTGAACATCGAGCCACA 4218
QY 4201 AAGTGTCTGTGTCTGTGATACCGTGCACCTTTGACCGCAGGCAATCCAAATGTCTCGAA 4260
Db 4219 AAGTGTCTGTGTCTGTGATACCGTGCACCTTTGACCGCAGGCAATCCAAATGTCTCGAA 4278
QY 4261 TGTCAAGGTGATGTGTCAACCCCAAGTGTCTCAAGTGTCTTGCAGCCACCTGGCGGTTCGCT 4320
Db 4279 TGTCAAGGTGATGTGTCAACCCCAAGTGTCTCAAGTGTCTTGCAGCCACCTGGCGGTTCGCT 4338
QY 4321 GGTGAATATGCCACACACTTTCACCGAGGCTTCTCCGTGTGACAAATGAATCCCCAGGT 4380
Db 4339 GGTGAATATGCCACACACTTTCACCGAGGCTTCTCCGTGTGACAAATGAATCCCCAGGT 4398
QY 4381 CTCAGAGCAAGAGAGCCAGACGAGCTTTCACCTGTGAAGGCTGTGATGAAGTGTGCCACG 4440
Db 4399 CTCAGAGCAAGAGAGCCAGACGAGCTTTCACCTGTGAAGGCTGTGATGAAGTGTGCCACG 4458
QY 4441 AATAACAAACGAGGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGATCAAAA 4500
Db 4459 AATAACAAACGAGGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGATCAAAA 4518
QY 4501 GTCCTCATTTATGACAAATGAAGCCAGAGAGCTGTGACAGAGCCGCTGGAGAGATTTGAG 4560
Db 4519 GTCCTCATTTATGACAAATGAAGCCAGAGAGCTGTGACAGAGCCGCTGGAGAGATTTGAG 4578
QY 4561 CTGTGCTTCCCGACGGGATGTATCTATTATGCTGCGCTTGTGCTTCCGAACTCGCA 4620
Db 4579 CTGTGCTTCCCGACGGGATGTATCTATTATGCTGCGCTTGTGCTTCCGAACTCGCA 4638
QY 4621 AATAAGCCAAAGCAGAGAAAAGCAGAGCTGTGATGCTTAACTGCTTGGAACTCCCTGCTG 4680
Db 4639 AATAAGCCAAAGCAGAGAAAAGCAGAGCTGTGATGCTTAACTGCTTGGAACTCCCTGCTG 4698
QY 4681 AAACCTGGAAGTGTAGACCGTGTAGACATGAAGTGTGACGCTGCTTCAAGTACACAGGTG 4740
Db 4699 AAACCTGGAAGTGTAGACCGTGTAGACATGAAGTGTGACGCTGCTTCAAGTACACAGGTG 4758
QY 4741 GTGTTGTGGGCAACCGAGGAAGGCTCTACGCGCTGAATGTCTTGAAGAACTCCCTAAC 4800
Db 4759 GTGTTGTGGGCAACCGAGGAAGGCTCTACGCGCTGAATGTCTTGAAGAACTCCCTAAC 4818
QY 4801 CATGTCCAGGAATGGAGCACTCTCCAAATTTATATTATCAAGGACTGGAGAGCTA 4860
Db 4819 CATGTCCAGGAATGGAGCACTCTCCAAATTTATATTATCAAGGACTGGAGAGCTA 4878
QY 4861 CTCATGATAGCAGGAGAGAGCGGCACTGTGCTTTGTGGACGTGAAGAAAGTGAACAG 4920
Db 4879 CTCATGATAGCAGGAGAGAGCGGCACTGTGCTTTGTGGACGTGAAGAAAGTGAACAG 4938
QY 4921 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCACCCCAATTTTGAAGCT 4980
Db 4939 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCACCCCAATTTTGAAGCT 4998
QY 4981 GTCAGGGCTGCCACTTTTGGGGCAGCAAGATTGAGAACGGGCTCTGCACTCTGTGCA 5040
Db 4999 GTCAAGGGCTGCCACTTTTGGGGCAGCAAGATTGAGAACGGGCTCTGCACTCTGTGCA 5058
QY 5041 GGCATGCCAGCAAGTGTCTATTCTCCGCTACACGAAACCTCAGCAAAATCTGCAATC 5100
Db 5059 GGCATGCCAGCAAGTGTCTATTCTCCGCTACACGAAACCTCAGCAAAATCTGCAATC 5118
QY 5101 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTCCACCAATTTACAGTATC 5160
Db 5119 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTCCACCAATTTACAGTATC 5178
QY 5161 CTCATTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTC 5220
Db 5179 CTCATTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTC 5238
QY 5221 CTGGATAAGAAATACCAATTCCTTGGGCACCTGCTGTGTTGGCCCTCTTCCACAGCTTC 5280

Db 5239 CTGGATAAGAAATACCAATTCCTTGGCACCTGTGTGTTTGGCCGCTTTCCTCAACAGCTTC 5298
QY 5281 CCTGTCTCAATCTGTGAGTGAACAGCGCAGGCGCAGGAGAGTACTTGTGTGTGTTTC 5340
Db 5299 CCTGTCTCAATCTGTGAGTGAACAGCGCAGGCGCAGGAGAGTACTTGTGTGTGTTTC 5358
QY 5341 CACGAATTTGGAGTGTTCGTGAGATTTCTTACGGAAGAGCTAGCCGCACAGACGATCTCAAG 5400
Db 5359 CACGAATTTGGAGTGTTCGTGAGATTTCTTACGGAAGAGCTAGCCGCACAGACGATCTCAAG 5418
QY 5401 TGGAGTCGCTTACCTTTGGCCTTTGGCTACAGAGAACCTTATCTGTGTTGTGACCCACTTC 5460
Db 5419 TGGAGTCGCTTACCTTTGGCCTTTGGCTACAGAGAACCTTATCTGTGTTGTGACCCACTTC 5478
QY 5461 AACTCACTCGAAGTAATTGAGATTCAGGCACTGCTCTCAGCAGGGACCTCTGCCCGAGCG 5520
Db 5479 AACTCACTCGAAGTAATTGAGATTCAGGCACTGCTCTCAGCAGGGACCTCTGCCCGAGCG 5538
QY 5521 TACCTGAGACATCCGACCCGCTACCTGGGCGCTTCCCTTCCCTCAGGAGCGATTTAC 5580
Db 5539 TACCTGAGACATCCGACCCGCTACCTGGGCGCTTCCCTTCCCTCAGGAGCGATTTAC 5598
QY 5581 TTGCGCTCTCATACACAGGATAAATTAAGGTCATTTGCTGCAAGGAAACCTCTGTGAAG 5640
Db 5599 TTGCGCTCTCATACACAGGATAAATTAAGGTCATTTGCTGCAAGGAAACCTCTGTGAAG 5658
QY 5641 GAGTCCGGCACTGAACACACCCGGGCGCTCACCTCCGCGCAGCAGCCCCACACAGCGA 5700
Db 5659 GAGTCCGGCACTGAACACACCCGGGCGCTCACCTCCGCGCAGCAGCCCCACACAGCGA 5718
QY 5701 GGCCACCCACGCTACACAGGACATCACCAAGCGGTGCGCTCCAGCCCGCGCGCCGCC 5760
Db 5719 GGCCACCCACGCTACACAGGACATCACCAAGCGGTGCGCTCCAGCCCGCGCGCCGCC 5778
QY 5761 GAAGGCCCGCAGCCACCCGCGAGAGCCAAAGCACACCCCAACCGCTACCGCGAGGGCGGACC 5820
Db 5779 GAAGGCCCGCAGCCACCCGCGAGAGCCAAAGCACACCCCAACCGCTACCGCGAGGGCGGACC 5838
QY 5821 GAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCTCGTGAGGAGAGAGTCCCGCGCGCGG 5880
Db 5839 GAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCTCGTGAGGAGAGAGTCCCGCGCGCGG 5898
QY 5881 ATGCTCAGCAGCGCGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGGGC 5940
Db 5999 ATGCTCAGCAGCGCGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGGGC 5958
QY 5941 CGGCTGCTGCGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAGGAGAGGGCAG 6000
Db 5959 CGGCTGCTGCGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAGGAGAGGGCAG 6018
QY 6001 AGTCCCTCTCAAGTTTTCACGTTTAACTGTACCTTATTTATGACTGGAATAAAAAAGCTG 6060
Db 6019 AGTCCCTCTCAAGTTTTCACGTTTAACTGTACCTTATTTATGACTGGAATAAAAAAGCTG 6078
QY 6061 GACAACTGCGCAGCTAACTGCTCAGGATCATCCAGCTGAATGAGAGAAATCCGG 6120
Db 6079 GACAACTGCGCAGCTAACTGCTCAGGATCATCCAGCTGAATGAGAGAAATCCGG 6138
QY 6121 CAGCAGGTGAAAGTCTGTTCTGAGAACAGATTATTGC 6159
Db 6139 CAGCAGGTGAAAGTCTGTTCTGAGAACAGATTATTGC 6177

RESULT 3
AX671108
LOCUS AX671108 8603 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 4 from Patent WO03004523.
ACCESSION AX671108
VERSION AX671108.1 GI:29329570
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

1921 CTCCAAGGAAACTCGAAGGCTG----- 1945
1921 CTCCAAGGAAACTCGAAGGCTGTAAGCCAGCAGGAGCCACCGAGCTGCTGCAG 1980
1946 -----CAAAAGAGCGAGCCGAGAGGGAGCTGGAGAGCTCGAAGCTCGAAGCCGAGAG 1992
1981 AATATCCGCCAGGCAAAAGGAGCGAGCCGAGAGGGAGCTGGAGAGCTCGAAGCCGAGAG 2040
1993 GATTCTTCTGAGGCGATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCGCCATTTCTCTG 2052
2041 GATTCTTCTGAGGCGATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCGCCATTTCTCTG 2100
2053 GAGAACAGAGTAAAGAGACTAGAGACCCTAGAGCCCTAGAGAAACAGACTGAGGATGAC 2112
2101 GAGAACAGGTAAGAGACTAGAGACCCTAGAGCCCTAGAGAAACAGACTGAGGATGAC 2160
2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2172
2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2220
2173 AAACATCGGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2232
2221 AAACATCGGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2280
2233 CAGCACTATGAGGAAAGAAATTAAGTTGTTGGACAATCAGATATAAGAAAGACCTGCTGAC 2292
2281 CAGCACTATGAGGAAAGAAATTAAGTTGTTGGACAATCAGATATAAGAAAGACCTGCTGAC 2340
2293 AAGGAGACTCGAGAACATGATGAGACACAGAGGAGGAGGCCCATGAGAGGCGCAA 2352
2341 AAGGAGACTCGAGAACATGATGAGACACAGAGGAGGAGGCCCATGAGAGGCGCAA 2400
2353 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA 2412
2401 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA 2460
2413 CAGAGATTGTGGAATGCTGAGGCCAATAAATCTGAGCAATAGCAGTCTTTTACC 2472
2461 CAGAGATTGTGGAATGCTGAGGCCAATAAATCTGAGCAATAGCAGTCTTTTACC 2520
2473 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
2521 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCGAAGCCGAAACTGGAGAGCAGCTGGAG 2592
2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCGAAGCCGAAACTGGAGAGCAGCTGGAG 2640
2593 AAGATCAGCCACCAAGACACACAGTGAACAAGATCGGCTGTAAGCTGGAGACCAAGATTG 2652
2641 AAGATCAGCCACCAAGACACACAGTGAACAAGATCGGCTGTAAGCTGGAGACCAAGATTG 2700
2653 CGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCAAGAG 2712
2701 CGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCAAGAG 2760
2713 CTAAGCTCTCCCTGGAGAGCCGAGTCAAGTTGACAGCCCTGAGGCTGACGGGG 2772
2761 CTAAGCTCTCCCTGGAGAGCCGAGTCAAGTTGACAGCCCTGAGGCTGACGGGG 2820
2773 GCCTCGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGCAAGCT 2832
2821 GCCTCGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGCAAGCT 2880
2833 GAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAATCAGCGCAAAATTTGATGCT 2892
2881 GAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAATCAGCGCAAAATTTGATGCT 2940
2893 CTTTCGTAAACAGCTGATCTGTAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 2952
2941 CTTTCGTAAACAGCTGATCTGTAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 3000
2953 GACACGCTGAATCAACCAAAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT 3012

3001 GACAAAGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT 3060
3013 GGGCCCAACACAGAGATTGTACAACCTGCGAAGTGAAGTGGACCATCTCCGCCGGAGATC 3072
3061 GGGCCCAACACAGAGATTGTACAACCTGCGAAGTGAAGTGGACCATCTCCGCCGGAGATC 3120
3073 ACGGAAACGAGAGATGCAAGCTTACAGCCAGAACAAACGATGGAGGCTCTGAAGACCAAG 3132
3121 ACGGAAACGAGAGATGCAAGCTTACAGCCAGAACAAACGATGGAGGCTCTGAAGACCAAG 3180
3133 TGCACCAATGCTGAGGAAACAGGTCATGGATTTCGAGGCCCTTAAACGATGAGCTGTAGAA 3192
3181 TGCACCAATGCTGAGGAAACAGGTCATGGATTTCGAGGCCCTTAAACGATGAGCTGTAGAA 3240
3193 AAAGAGCGGAGCTGGGAGGCTGAGGAGGCTCTGGGTGATGAGAAATCCCAAGTTTGAG 3252
3241 AAAGAGCGGAGCTGGGAGGCTGAGGAGGCTCTGGGTGATGAGAAATCCCAAGTTTGAG 3300
3253 TGTCCGGTTCCAGAGCTGCGAGAAATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3312
3301 TGTCCGGTTCCAGAGCTGCGAGAAATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3360
3313 GATCAGCGGATCACCGAGTCTCCGAGGTGGAGCTGCGAGTGAAGGAGACCAAGGCT 3372
3361 GATCAGCGGATCACCGAGTCTCCGAGGTGGAGCTGCGAGTGAAGGAGACCAAGGCT 3420
3373 GAGATTCTCCGCTCTGACAGGCTCTCAAAAGACAGAAAGTGAAGCCGAGAGCTCTCT 3432
3421 GAGATTCTCCGCTCTGACAGGCTCTCAAAAGACAGAAAGTGAAGCCGAGAGCTCTCT 3480
3433 GACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGATGATGCTCCGAGGTTA 3492
3481 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGATGATGCTCCGAGGTTA 3540
3493 CAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAA 3552
3541 CAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAA 3600
3553 TTAACAGAGCAGATGGAACCTGCAAGAAAATCAATTTCCGCTGATCTGAAGCTGCA 3612
3601 TTAACAGAGCAGATGGAACCTGCAAGAAAATCAATTTCCGCTGATCTGAAGCTGCA 3660
3613 GAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3672
3661 GAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3720
3673 GAAACCAATTCAGGTTCTTATTTCTATGAAAGGTGAAAATGGAAGGCACTATTTCTCAA 3732
3721 GAAACCAATTCAGGTTCTTATTTCTATGAAAGGTGAAAATGGAAGGCACTATTTCTCAA 3780
3733 CAAACCAACTCATTTGATTTTCTGAGCCAAATGCAACCTGCTTAAAGAGAAAAG 3792
3781 CAAACCAACTCATTTGATTTTCTGAGCCAAATGCAACCTGCTTAAAGAGAAAAG -- 3838
3793 GGTATTATTAGTCGACGAAAGAGGACCTGCTTTTACCACACAGGTTCTCTCAGTAC 3852
3839 -----AGGTTCTCTCAGTAC 3855
3853 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAAAGCTGCTGTGAGAGCTAGAGAGGAC 3912
3856 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAAAGCTGCTGTGAGAGCTAGAGAGGAC 3915
3913 CTTTCAGAGACCCGATCGAGCTCCGGTCCGCGGAGGAGAGCTGCCACCGCAAGCA 3972
3916 CTTTCAGAGACCCGATCGAGCTCCGGTCCGCGGAGGAGAGCTGCCACCGCAAGCA 3975
3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATGCTCCGCC 4032
3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATGCTCCGCC 4035
4033 ATCTGGGCTGCGCAGAGCACCGCCAGTGCATGAGCTGCTGGCCCGCCATCCAGC 4092

Db 4036 ATCGTGGGTGCGCAGAGCACCAGCCAGTGCCATGAGCCTGCTGGGCCCCGCCATCCAGC 4095
Qy 4093 CGCAGAAAGAGTCTTCAATCCAGAGGAATTTAGTCGGCGTCTTAAAGGAACCCATGCAC 4152
Db 4096 CGCAGAAAGAGTCTTCAATCCAGAGGAATTTAGTCGGCGTCTTAAAGGAACCCATGCAC 4155
Qy 4153 CACAATATTCCTCACCAGATCAACGTAAGGACTGAACATGAGCCAGCCAAAGTGTGCTGTG 4212
Db 4156 CACAATATTCCTCACCAGATCAACGTAAGGACTGAACATGAGCCAGCCAAAGTGTGCTGTG 4215
Qy 4213 TGTCTGATACCGTGCACTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCAAGTGTGATG 4272
Db 4216 TGTCTGATACCGTGCACTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCAAGTGTGATG 4275
Qy 4273 TGTCAACCCCAAGTGTCCAGTGTCTGGCAGCCACCTGGCGCTTGGCTGTGTAATATGCC 4332
Db 4276 TGTCAACCCCAAGTGTCCAGTGTCTGGCAGCCACCTGGCGCTTGGCTGTGTAATATGCC 4335
Qy 4333 ACACACTTCACCGAGGCTTTCTGCGGTGACAAATGAATCACTCCCGAGGTCTCCAGACCAAG 4392
Db 4336 ACACACTTCACCGAGGCTTTCTGCGGTGACAAATGAATCACTCCCGAGGTCTCCAGACCAAG 4395
Qy 4393 GAGCCAGCAGCAGCTTGCACTTGAAGGTGATGAGTGCCTGAGGATGAGGATGAACCAAGCA 4452
Db 4396 GAGCCAGCAGCAGCTTGCACTTGAAGGTGATGAGTGCCTGAGGATGAGGATGAACCAAGCA 4455
Qy 4453 GGACAGCAAGGCTGGGACAGGAATACATTTCTGTGAGGATCAAAAGTCTCTCATTTAT 4512
Db 4456 GGACAGCAAGGCTGGGACAGGAATACATTTCTGTGAGGATCAAAAGTCTCTCATTTAT 4515
Qy 4513 GACAAATGAAGCCAGAGAGCTGACAGAGCCCGGTGGAAGAAATTTGAGTGTGCTTCCC 4572
Db 4516 GACAAATGAAGCCAGAGAGCTGACAGAGCCCGGTGGAAGAAATTTGAGTGTGCTTCCC 4575
Qy 4573 GACGGGATGTACTATTATGTGTCGCTTGGTGTCTTCCGAACTGCAATACAGCCAA 4632
Db 4576 GACGGGATGTACTATTATGTGTCGCTTGGTGTCTTCCGAACTGCAATACAGCCAA 4635
Qy 4633 GCA----- 4635
Db 4636 GCAGATGTCCATACATACCTGAGATCTCACCCGACACACCTGCTGGCCCGG 4695
Qy 4636 ----- 4635
Db 4696 AGAACCTCTACTTGTAGCTCCAGTCCAGTCCCTGTGACAAACAGCGTGGGTCCACGCCCTTA 4755
Qy 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662
Db 4756 GAATCAGTTGTGCGAGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAACTG 4815
Qy 4663 CTTGGAAACTCCCTGTGTAACCTGGAAGTGTATGACCTGTAGACATGAACTGCAAGCTG 4722
Db 4816 CTTGGAAACTCCCTGTGTAACCTGGAAGTGTATGACCTGTAGACATGAACTGCAAGCTG 4875
Qy 4723 CCCTTCAGTGACAGAGTGTGTTGGTGGCCACCGAGGAAGGCTCTACGCCCTGAAATGTC 4782
Db 4876 CCCTTCAGTGACAGAGTGTGTTGGTGGCCACCGAGGAAGGCTCTACGCCCTGAAATGTC 4935
Qy 4783 TTGAAAAATCCCTTAACCATGTCACAGAAATGGAGCAGTCTTCCAAATTTATATTATC 4842
Db 4936 TTGAAAAATCCCTTAACCATGTCACAGAAATGGAGCAGTCTTCCAAATTTATATTATC 4995
Qy 4843 AAGGACCTGGAGAGCTACTCATGTATAGCAGAGAGAGCGGCACTGTGCTTTGGAC 4902
Db 4996 AAGGACCTGGAGAGCTACTCATGTATAGCAGAGAGAGCGGCACTGTGCTTTGGAC 5055
Qy 4903 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 4962
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 5115
Qy 4963 CCGACATTTTGAAGCTGTCAAGGCTGCACTTTGTTGGGAGGCAAGATTCAGAAC 5022
Db 5116 CCGACATTTTGAAGCTGTCAAGGCTGCACTTTGTTGGGAGGCAAGATTCAGAAC 5175

Qy 5023 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTGTCATTTCTCCGTACAAACGAAAC 5082
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTGTCATTTCTCCGTACAAACGAAAC 5235
Qy 5083 CTCACAAATACTGCATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5142
Db 5236 CTCACAAATACTGCATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
Qy 5143 TTACCAATACAGTATCTCTCAITGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5202
Db 5296 TTACCAATACAGTATCTCTCAITGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
Qy 5203 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCAATTCCTTGGCACTGTGTGTTGCC 5262
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCAATTCCTTGGCACTGTGTGTTGCC 5415
Qy 5263 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGCAGAG 5322
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGCAGAG 5475
Qy 5323 GAGTACTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATCTTACGGAAGACGTAGC 5382
Db 5476 GAGTACTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATCTTACGGAAGACGTAGC 5535
Qy 5383 CGCACAGACGATCTCAAGTGGAGTGTCTTACCTTTGGCTTTGGCTACAGAGAACCTAT 5442
Db 5536 CGCACAGACGATCTCAAGTGGAGTGTCTTACCTTTGGCTTTGGCTACAGAGAACCTAT 5595
Qy 5443 CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATTTAGGGTCAATTTGCTGC 5622
Db 5596 CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATTTAGGGTCAATTTGCTGC 5775
Qy 5503 GGGACCTCTGCGCAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAAT 5562
Db 5556 GGGACCTCTGCGCAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAAT 5715
Qy 5563 TCCTCAGAGCGATTTACTTTGGGTCTCTATACAGGATAAATTTAAGGGTCAATTTGCTGC 5622
Db 5716 TCCTCAGAGCGATTTACTTTGGGTCTCTATACAGGATAAATTTAAGGGTCAATTTGCTGC 5775
Qy 5623 AAGGAAACCTGTGAGGAGTCCGCACTGAAACACCGCGGCGCTCAGCTCCCGC 5682
Db 5776 AAGGAAACCTGTGAGGAGTCCGCACTGAAACACCGCGGCGCTCAGCTCCCGC 5835
Qy 5683 AGCAGCCCCAACAGCGAGGCCACCACTGACAAAGCAGACATCAAGCGCGTGGCC 5742
Db 5836 AGCAGCCCCAACAGCGAGGCCACCACTGACAAAGCAGACATCAAGCGCGTGGCC 5895
Qy 5743 TCAGACCCAGCGCGCGCGAGGCCACCACTGAGAGAGCAAGCAACCCCAAGC 5802
Db 5896 TCAGACCCAGCGCGCGCGAGGCCACCACTGAGAGAGCAAGCAACCCCAAGC 5955
Qy 5803 TACCGCAGGGCGGAGCCGAGCTGCGCAGGACAAAGTCTCTGCGCGCGCTTGGAGCGA 5862
Db 5956 TACCGCAGGGCGGAGCCGAGCTGCGCAGGAGCAAGTCTCTGCGCGCGCTTGGAGCGA 6015
Qy 5863 GAGAAGTCCCGCGCGGATGTCTAGCAACCGAGAGAGCGGTCCCGCGGAGCGTGT 5922
Db 6016 GAGAAGTCCCGCGCGGATGTCTAGCAACCGAGAGAGCGGTCCCGCGGAGCGTGT 6075
Qy 5923 GAGACAGCAGAGGCGCGGTCTGCGGAGCGCTGAGGACCCGCTGTCCAGGTTG 5982
Db 6076 GAGACAGCAGAGGCGCGGTCTGCGGAGCGCTGAGGACCCGCTGTCCAGGTTG 6135
Qy 5983 AACAGGAAAGGCGAGGTGCTCTCAAGTTTTCACGGTTTAACTGTCACTTAT 6042
Db 6136 AACAGGTTCTGGACAGTCTTCAAGTATTAATCTCAGCAGAAAAACCACTCTCATCT 6195

RESULT 4
AX671112
LOCUS

AX671112 6156 bp DNA linear PAT 27-MAR-2003

DEFINITION Sequence 8 from Patent WO03004523.
ACCESSION AX671112
VERSION AX671112.1 GI:29329572
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Zhu, Z.
TITLE Regulation of human citron rho/rac-interacting kinase
JOURNAL Patent: WO 03004523-A 8 16-JAN-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..6156
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 92.0%; Score 5666.6; DB 6; Length 6156;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 5945; Conservative 0; Mismatches 9; Indels 243; Gaps 3;
QY 1 ATGTTGAAGTTCAAATATATGAGCGCGGAATCCCTTTGGATGCTGGTGTGCTGAACCCATT 60
DB 1 ATGTTGAAGTTCAAATATATGAGCGCGGAATCCCTTTGGATGCTGGTGTGCTGAACCCATT 60
QY 61 GCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 120
DB 61 GCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 120
QY 121 CAGCAGATGTCCTCTTTCCGAGAGGATATATAGATGCCCTCTTTGTTCTTTTGA 180
DB 121 CAGCAGATGTCCTCTTTCCGAGAGGATATATAGATGCCCTCTTTGTTCTTTTGA 180
QY 181 GAATCGAGTCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 GAATCGAGTCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAACTCAGA 300
DB 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAACTCAGA 300
QY 301 AGTCTTGTAGTGTGGTCACTTTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 360
DB 301 AGTCTTGTAGTGTGGTCACTTTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 360
QY 361 GACATCTATGCTATGAAAGTATGAAAGAGAGGCTTTATTGCGCCAGGAGCAGGTTTCA 420
DB 361 GACATCTATGCTATGAAAGTATGAAAGAGAGGCTTTATTGCGCCAGGAGCAGGTTTCA 420
QY 421 TTTTGTGAGGAGAGCGGAACATAATATCTCGAAGCACAAGCCCGTGGATCCCCAATTA 480
DB 421 TTTTGTGAGGAGAGCGGAACATAATATCTCGAAGCACAAGCCCGTGGATCCCCAATTA 480
QY 481 CAGTATGCCCTTTCCAGGACAAAATCACCTTTATCTGATGAGGATATCAGCTCGAGGG 540
DB 481 CAGTATGCCCTTTCCAGGACAAAATCACCTTTATCTGATGAGGATATCAGCTCGAGGG 540
QY 541 GACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600
DB 541 GACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCTCTGATGGATACGTCATCGA 660
DB 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCTCTGATGGATACGTCATCGA 660
QY 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCAGCAGGACATCAAGCTGGTGAATTT 720
DB 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCAGCAGGACATCAAGCTGGTGAATTT 720
QY 721 GGATCTGCCCGGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780

721 GGATCTGCCCGGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
QY 781 CCAGATTACATGGCTCCTGAAGTCTGACTGTGATGAACGGGATGAGAAAGCAGCTAC 840
DB 781 CCAGATTACATGGCTCCTGAAGTCTGACTGTGATGAACGGGATGAGAAAGCAGCTAC 840
QY 841 GGCCTGACATGTGACTGGTGTGAGTGGGCGTGAATTCCTATGAGATGATTTATGGGAGA 900
DB 841 GGCCTGACATGTGACTGGTGTGAGTGGGCGTGAATTCCTATGAGATGATTTATGGGAGA 900
QY 901 TCCCTTTCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 960
DB 901 TCCCTTTCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 960
QY 961 TTTTGAATTTTCAGATGACCCCAAGTGAAGTGAAGTCTTTGCTGCCATCTTTCTTC 1020
DB 961 TTTTGAATTTTCAGATGACCCCAAGTGAAGTGAAGTCTTTGCTGCCATCTTTCTTC 1020
QY 1021 TTGTTGCGGCCAGAAAGAGAGACTGAATTTGAAGTCTTTGCTGCCATCTTTCTTC 1080
DB 1021 TTGTTGCGGCCAGAAAGAGAGACTGAATTTGAAGTCTTTGCTGCCATCTTTCTTC 1080
QY 1081 TCTAAATTTGACTGGACAAACATTCGTAACCTCTCCCTCCCTTCGTTCCACCTCAAG 1140
DB 1081 TCTAAATTTGACTGGACAAACATTCGTAACCTCTCTCCCTCTCGTTCCACCTCAAG 1140
QY 1141 TCTGACATGACACCTCCAAATTTTGAATGAACAGAGAGAAATTCGTGGGTTTCACTCT 1200
DB 1141 TCTGACATGACACCTCCAAATTTTGAATGAACAGAGAGAAATTCGTGGGTTTCACTCT 1200
QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGGGGTTTTCG 1260
DB 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGATTTCTTGTGATGATCTGAGTCTGTGTCGGGCTCGGACTCC 1320
DB 1261 TACAGCAAGGCACTGGGATTTCTTGTGATGATCTGAGTCTGTGTCGGGCTCGGACTCC 1320
QY 1321 CCGTCCAGGCACTAGCTCCATGGAAAAAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
DB 1321 CCGTCCAGGCACTAGCTCCATGGAAAAAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
QY 1381 TCTCAGCAAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCGAGAGTGTCA 1440
DB 1381 TCTCAGCAAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCGAGAGTGTCA 1440
QY 1441 GAGGTGAGGCTGTGCTTAGTTCAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
DB 1441 GAGGTGAGGCTGTGCTTAGTTCAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
QY 1501 TCCCTCTGAGCAGGACCTTGTCTACCTACATCAGAAATGCAGTAGCTTAAAGCGAAGT 1560
DB 1501 TCCCTCTGAGCAGGACCTTGTCTACCTACATCAGAAATGCAGTAGCTTAAAGCGAAGT 1560
QY 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1620
DB 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1620
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACCAGGCTCAA 1680
DB 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACCAGGCTCAA 1680
QY 1681 GTGCAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740
DB 1681 GTGCAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740
QY 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTGGGCTTGTCTGAGAAATTTCAAG 1800
DB 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTGGGCTTGTCTGAGAAATTTCAAG 1800
QY 1801 CGGAAACGACAGATGTCAAGTAACTTGTGAAGCTAAGGATCAAGGAGAGCTGAA 1860

Db	1801	CGGAAAGCGACAGAAATGTCAGCATAAACTGTTGAAGGCTAAGGATCAAGGGAAGCCGTGAA	1861
Qy	1861	GTGGGAAATATCGAAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1920
Db	1861	GTGGGAAATATCGAAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1920
Qy	1921	CTCCAGAGAAACTGGAGAGGCT-----	1944
Db	1921	CTCCAGAGAAACTGGAGAGGCTGTAAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAG	1980
Qy	1945	-----GCAAAGGACGAGCCGAGAGGGAGCTGCAGAAAGCTGCAGAAACCCGAGAG	1992
Db	1981	AAATATCCGCCAGGCAAGGAGCGAGCCGAGAGGGAGCTGGAGAAAGCTGCAGAAACCCGAGAG	2040
Qy	1993	GATTTCTTTGAAGGCATCAGAAAGAAAGCTGGTGGAAAGCTCAGGAAGCCGCCCAATCTCTG	2052
Db	2041	GATTTCTTTGAAGGCATCAGAAAGAAAGCTGGTGGAAAGCTGAGGAAAGCCGCCCAATCTCTG	2100
Qy	2053	GAGAACAGGTTAAGAGACTAGACACATGGAGCGTAGAGAAACAGACATGAAAGATGAC	2112
Db	2101	GAGAACAGGTTAAGAGACTAGACACATGGAGCGTAGAGAAACAGACATGAAAGATGAC	2160
Qy	2113	ATCCAGACAAAATCCCACACAGATCCAGCAGATGGCTGATAAAATTCTGGAGCTCGAAGAG	2172
Db	2161	ATCCAGACAAAATCCCACACAGATCCAGCAGATGGCTGATAAAATTCTGGAGCTCGAAGAG	2220
Qy	2173	AAACATCGGAGGCCCAAGTCTCAGCCACGACCTAGAAAGTGCACCTGAAACAGAAAGAG	2232
Db	2221	AAACATCGGAGGCCCAAGTCTCAGCCACGACCTAGAAAGTGCACCTGAAACAGAAAGAG	2280
Qy	2233	CAGCACTATCAGGAAAAAGATTAAAGTGTTCGACAAATCAGATAAAGAAAGACCTGGCTGAC	2292
Db	2281	CAGCACTATCAGGAAAAAGATTAAAGTGTTCGACAAATCAGATAAAGAAAGACCTGGCTGAC	2340
Qy	2293	AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAAGGCGAAA	2352
Db	2341	AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAAGGCGAAA	2400
Qy	2353	ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA	2412
Db	2401	ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA	2460
Qy	2413	CAGAGGATTGTGAACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGCACTTTTTTACC	2472
Db	2461	CAGAGGATTGTGAACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGCACTTTTTTACC	2520
Qy	2473	CAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC	2532
Db	2521	CAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC	2580
Qy	2533	CTGGAGACACAGGCTGGGAAGTTGGAGGCCCGAGAAACCTGGAGGAGCAGCTGGAG	2592
Db	2581	CTGGAGACACAGGCTGGGAAGTTGGAGGCCCGAGAAACCTGGAGGAGCAGCTGGAG	2640
Qy	2593	AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTGCTGAACTGGAGACCAAGATTG	2652
Db	2641	AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTGCTGAACTGGAGACCAAGATTG	2700
Qy	2653	CGGAGGTCAGTCTTAGAGCACGAGGAGCAAAACTGGAGCTCAAGCGCCAGCTCACAAG	2712
Db	2701	CGGAGGTCAGTCTTAGAGCACGAGGAGCAAAACTGGAGCTCAAGCGCCAGCTCACAAG	2760
Qy	2713	CTACAGCTCTCCCTGCAAGGAGCGAGTCAAGTTGACAGCCCTGCAAGGCTGACCGGGCG	2772
Db	2761	CTACAGCTCTCCCTGCAAGGAGCGAGTCAAGTTGACAGCCCTGCAAGGCTGACCGGGCG	2820
Qy	2773	GCCTTGAGAGGCAGCTTCCGAGGCGAAGACAGAGCTGGAAGAGACCAACAGCAGAAGCT	2832
Db	2821	GCCTTGAGAGGCAGCTTCCGAGGCGAAGACAGAGCTGGAAGAGACCAACAGCAGAAGCT	2880
Qy	2833	GAAGAGAGATCCAGGCACCTCACGGCACATAGAGATGAAATCCAGCGCAAAATTGATGCT	2892
Db	2881	GAAGAGAGATCCAGGCACCTCACGGCACATAGAGATGAAATCCAGCGCAAAATTGATGCT	2940

QY	2893	CTTCGTAACGAGCTGTACTGTAATCACAGACCTGGAGGAGCAGTTAAACCAAGCTGACCGAG	2952
DB	2941	CTTCGTAACGAGCTGTACTGTAATCACAGACCTGGAGGAGCAGTTAAACCAAGCTGACCGAG	3000
QY	2953	GACAAACGCTGAACCTCAACAACCAAACTTCTACTTCTCCAAAACAATCCGATGAGGCTTCT	3012
DB	3001	GACAAACGCTGAACCTCAACAACCAAACTTCTACTTCTCCAAAACAATCCGATGAGGCTTCT	3060
QY	3013	GGCGCCAAACGACGAGATTGTACAACCTCGAAGTCAAGTGCAGACCATCTCCGCCCGGAGATC	3072
DB	3061	GGCGCCAAACGACGAGATTGTACAACCTCGAAGTCAAGTGCAGACCATCTCCGCCCGGAGATC	3120
QY	3073	ACGGAAACGAGAGATGAGCTTACGAGCCGAAACCAACGATGAGAGCTCTGAAGACCAAGC	3132
DB	3121	ACGGAAACGAGAGATGAGCTTACGAGCCGAAACCAACGATGAGAGCTCTGAAGACCAAGC	3180
QY	3133	TGCACCAATGCTGGAGGAACAGGTCATGGATTTCGGAGGCCCTAAACCATGAGCTGCTAGAA	3192
DB	3181	TGCACCAATGCTGGAGGAACAGGTCATGGATTTCGGAGGCCCTAAACCATGAGCTGCTAGAA	3240
QY	3193	AAAGAGCGGCAATGGGAGGCTCTGGAGAGCGTCTCTGGGTGATGAGAAATCCCAAGTTTGAG	3252
DB	3241	AAAGAGCGGCAATGGGAGGCTCTGGAGAGCGTCTCTGGGTGATGAGAAATCCCAAGTTTGAG	3300
QY	3253	TGTCGGGTTCGAGAGCTGACAGAAATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCC	3312
DB	3301	TGTCGGGTTCGAGAGCTGACAGAAATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCC	3360
QY	3313	GATCAGCGGATCACCGAGTCTCCGCCAGGTGGTGAGCTGGCAGTGAAGGAGCAACAAGGT	3372
DB	3361	GATCAGCGGATCACCGAGTCTCCGCCAGGTGGTGAGCTGGCAGTGAAGGAGCAACAAGGT	3420
QY	3373	GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAACTGAGGCCGAGAGCCCTCTCT	3432
DB	3421	GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAACTGAGGCCGAGAGCCCTCTCT	3480
QY	3433	GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTTGAATGAATCCCGAAGCTTA	3492
DB	3481	GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTTGAATGAATCCCGAAGCTTA	3540
QY	3493	CAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGTTCTGGAAGCAGACGCCAAA	3552
DB	3541	CAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGTTCTGGAAGCAGACGCCAAA	3600
QY	3553	TTACAGCAGCAGATGACCTGCAGAAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA	3612
DB	3601	TTACAGCAGCAGATGACCTGCAGAAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA	3660
QY	3613	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTACTTGGAGTATCAGCTG	3672
DB	3661	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTACTTGGAGTATCAGCTG	3720
QY	3673	GAAACATTCAGGTTCTTATTTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAA	3732
DB	3721	GAAACATTCAGGTTCTTATTTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAA	3780
QY	3733	CAAAACCAACTCATTTGATTTCTGCAAGCCAAATGGACCAACCTGCTAAAAAGAAAAAG	3792
DB	3781	CAAAACCAACTCATTTGATTTCTGCAAGCCAAATGGACCAACCTGCTAAAAAGAAAA--	3838
QY	3793	GGTTTATTAGTCGACGGAAAGAGGCCCTGCTTTACCCACACAGGTTCTCTGCAAGTAC	3852
DB	3839	-----AGGTTTCTCTGCAAGTAC	3855
QY	3853	AATGAGCTGAAGCTGGCCCTGGAGAGAGGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC	3912
DB	3856	AATGAGCTGAAGCTGGCCCTGGAGAGAGGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC	3915
QY	3913	CTTCAGAAGACCCGATCGAGCTCCGGTCCGCCCGGAGGAAAGCTGCCACCCGCAAGCA	3972
DB	3916	CTTCAGAAGACCCGATCGAGCTCCGGTCCGCCCGGAGGAAAGCTGCCACCCGCAAGCA	3975

QY 3973 ACGGACCCACACCCATCCAGCCAGCCACCGCGAGGACGAGATCGCCATGTCGGCC 4032
Db 3976 ACGGACCCACACCCATCCAGCCAGCCACCGCGAGGACGAGATCGCCATGTCGGCC 4035
QY 4033 ATCGTCGGGTCCGAGACGACGACCCAGTCGATGAGCCTGCTGCCCGCCATCCAGC 4092
Db 4036 ATCGTCGGGTCCGAGACGACGACCCAGTCGATGAGCCTGCTGCCCGCCATCCAGC 4095
QY 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCAATGCAC 4152
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCAATGCAC 4155
QY 4153 CACAATATTCCTCACCGATTCAGCTAGACTGAAATCGAGGACCAACAAGTGTGCTGTG 4212
Db 4156 CACAATATTCCTCACCGATTCAGCTAGACTGAAATCGAGGACCAACAAGTGTGCTGTG 4215
QY 4213 TGTCTGGATACCGTGCACCTTGGAGCCAGGATCCAAATGCTCGAATGTCAGGTGATG 4272
Db 4216 TGTCTGGATACCGTGCACCTTGGAGCCAGGATCCAAATGCTCGAATGTCAGGTGATG 4275
QY 4273 TGTACCCCAAGTGTCCAGCTGCTGTCAGCCACCTGCGGTTGCTGCTGAATATGCC 4332
Db 4276 TGTACCCCAAGTGTCCAGCTGCTGTCAGCCACCTGCGGTTGCTGCTGAATATGCC 4335
QY 4333 ACACACTTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCGAGGTCCTCCAGACCAAG 4392
Db 4336 ACACACTTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCGAGGTCCTCCAGACCAAG 4395
QY 4393 GAGCCGACGAGCTTGACCTTGAAGGAGGTGGATGAAGGTCGCCAGGAATAACAAACGA 4452
Db 4396 GAGCCGACGAGCTTGACCTTGAAGGAGGTGGATGAAGGTCGCCAGGAATAACAAACGA 4455
QY 4453 GGACAGAGGTCGGACAGGAAGTACATGTCCTGGAGGATCAAAAGTCTCTATTAT 4512
Db 4456 GGACAGAGGTCGGACAGGAAGTACATGTCCTGGAGGATCAAAAGTCTCTATTAT 4515
QY 4513 GACAAATGAAGCAGAGAACTGGAAGGCGGTGGAAATTTGAGTGTGCTTCCC 4572
Db 4516 GACAAATGAAGCAGAGAACTGGAAGGCGGTGGAAATTTGAGTGTGCTTCCC 4575
QY 4573 GACGGGATGTAATCTATTCATGGTCCGTTGTTGTTGCTGCGAACTCGCAAAATACGCCAAA 4632
Db 4576 GACGGGATGTAATCTATTCATGGTCCGTTGTTGTTGCTGCGAACTCGCAAAATACGCCAAA 4635
QY 4633 GCA----- 4635
Db 4636 GCAGATGTCCTCATACATCTGAAGATGGAATCTACCCGACACCACTGCTGGCCCGGG 4695
QY 4636 ----- 4635
Db 4696 AGAAACCTCTACTTGTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTACACGCCCTTA 4755
QY 4636 ----- 4635
Db 4756 GAATCAGTTGTGCGAGGTGGAGAGTTTCTAGGGGAAAAGCAGAAAGCTGATGCTAAACTG 4815
QY 4663 CTTGGAACCTCCCTGCTGAAACTGGAAGGTGATGACCGGTCTAGACATGAACCTGACGCTG 4722
Db 4816 CTTGGAACCTCCCTGCTGAAACTGGAAGGTGATGACCGGTCTAGACATGAACCTGACGCTG 4875
QY 4723 CCGTTCAGTGAACAGTGTGTTGTTGGGCAACGAGGAAGGCTCTACGCGCTGAATGTC 4782
Db 4876 CCGTTCAGTGAACAGTGTGTTGTTGGGCAACGAGGAAGGCTCTACGCGCTGAATGTC 4935
QY 4783 TTGAAAACCTCCCTAACCCATCTCCAGGAATGGAGGAGTCTCCAAATTTATATTATC 4842
Db 4936 TTGAAAACCTCCCTAACCCATCTCCAGGAATGGAGGAGTCTCCAAATTTATATTATC 4995
QY 4843 AAGGACCTGGAGAGCTACTCATATGATAGCAGGAGAGGCGGCACTGTGCTTGTGGAC 4902
Db 4996 AAGGACCTGGAGAGCTACTCATATGATAGCAGGAGAGGCGGCACTGTGCTTGTGGAC 5055
QY 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCACCGCCGACATCTCA 4962

Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5115
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTGTTTGGGGCAGGCAAGATTGAGAAC 5022
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTGTTTGGGGCAGGCAAGATTGAGAAC 5175
QY 5023 GGGCTCTGCTCTGTGAGGCTATCCCGAGCAAAAGTCTGCTCATTTCTCCGCTACAAACGAAAAC 5082
Db 5176 GGGCTCTGCTCTGTGAGGCTATCCCGAGCAAAAGTCTGCTCATTTCTCCGCTACAAACGAAAAC 5235
QY 5083 CTCAGCAAAATACTGCTCATCCGGAAGAGATAGAGACTCTCAGAGCCCTGAGCTGTATCCAC 5142
Db 5236 CTCAGCAAAATACTGCTCATCCGGAAGAGATAGAGACTCTCAGAGCCCTGAGCTGTATCCAC 5295
QY 5143 TTACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTCTACGAAATCGCATGAGCAG 5202
Db 5296 TTACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTCTACGAAATCGCATGAGCAG 5355
QY 5203 TACAGCTTCGAGGAATTTCTGGATAAGAAATGAACATTTCTTTGGCACTGCTGTGTTGCC 5262
Db 5356 TACAGCTTCGAGGAATTTCTGGATAAGAAATGAACATTTCTTTGGCACTGCTGTGTTGCC 5415
QY 5263 GCCTCTTCCACAGCTTCCCTGCTCAATCGTCGAGGTGAACAGCGCAGGCGACGAGAG 5322
Db 5416 GCCTCTTCCACAGCTTCCCTGCTCAATCGTCGAGGTGAACAGCGCAGGCGACGAGAG 5475
QY 5323 GAGTACTTGTGTTTCCACGAAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5382
Db 5476 GAGTACTTGTGTTTCCACGAAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5535
QY 5383 CGCACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGGCTACAGAGAACCTTAT 5442
Db 5536 CGCACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGGCTACAGAGAACCTTAT 5595
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGACGCTCTCTCAGCA 5502
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGACGCTCTCTCAGCA 5655
QY 5503 GGGACCCCTGCGGAGCGTACCTGGAATCCCGAAACCGCGCTACCTGGGCGCTGCCATT 5562
Db 5656 GGGACCCCTGCGGAGCGTACCTGGAATCCCGAAACCGCGCTACCTGGGCGCTGCCATT 5715
QY 5563 TCCTCAGAGCGATTTACTTGGGCTCTCATACAGGATAAATTAAGGTCATTTGCTGC 5622
Db 5716 TCCTCAGAGCGATTTACTTGGGCTCTCATACAGGATAAATTAAGGTCATTTGCTGC 5775
QY 5623 AAGGGAACCTCTGTGAAGAGTCCGGCACTGAACACCAACCGGGCGCTCCACCTCCCGC 5682
Db 5776 AAGGGAACCTCTGTGAAGAGTCCGGCACTGAACACCAACCGGGCGCTCCACCTCCCGC 5835
QY 5683 AGCAGCCCGAACAGCGAGGCGCCACCACTGACAGGACATCACCAGCGGTGGCC 5742
Db 5836 AGCAGCCCGAACAGCGAGGCGCCACCACTGACAGGACATCACCAGCGGTGGCC 5895
QY 5743 TCCAGCCAGCGCGCCGAGGCGCCAGCACTCCGAGAGCCAGCACACCCACCGC 5802
Db 5896 TCCAGCCAGCGCGCCGAGGCGCCAGCACTCCGAGAGCCAGCACACCCACCGC 5955
QY 5803 TACCGCAGGCGGCGGACCGGAGTGTGCGAGGGAACAAGTCTCTGCGCGCCCTTGGAGCGA 5862
Db 5956 TACCGCAGGCGGCGGACCGGAGTGTGCGAGGGAACAAGTCTCTGCGCGCCCTTGGAGCGA 6015
QY 5863 GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGGGAGGCTGTTT 5922
Db 6016 GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGGGAGGCTGTTT 6075
QY 5923 GAAGACAGCAGCGGCGGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCCAGGTG 5982
Db 6076 GAAGACAGCAGCGGCGGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCCAGGTG 6135
QY 5983 AACAGGAGAGAGGGA 5999

Db	6136 AACAGGTGAGCGACA 6152	Db	595 GACTTGTGTCACTTTTGAATAGATATAGGACGAGTTAGATGAAACCTGTACACAGTTT 654
RESULT 5		Qy	601 TACCTAGCTGAGTGAATTTTGGCTGTTTACAGCGTTTCATCTGTATGGGATACGTGCAATCGA 660
LOCUS	AX504254	Db	655 TACCTAGCTGAGTGAATTTTGGCTGTTTACAGCGTTTCATCTGTATGGGATACGTGCAATCGA 714
DEFINITION	Sequence 43 from Patent WO0233099.	Qy	661 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
ACCESSION	AX504254	Db	715 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 774
VERSION	AX504254.1 GI:23386094	Qy	721 GGATCTCCGCGAAATGAATCAAAACAGATGCTGAATGCCAAACTCCCGATTGGGACC 780
KEYWORDS	Homo sapiens (human)	Db	775 GGATCTCCGCGAAATGAATCAAAACAGATGCTGAATGCCAAACTCCCGATTGGGACC 834
SOURCE	Homo sapiens	Qy	781 CCAGATTACATGGCTCTGAAGTGTGACTGTATGAAACCGGGATGAAAGGACCTTAC 840
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Db	835 CCAGATTACATGGCTCTGAAGTGTGACTGTATGAAACCGGGATGAAAGGACCTTAC 894
REFERENCE	1 Gururajan, R., Baughn, M.R., Wallia, N.K., Elliott, V.S., Xu, Y., Arvizu, C., Yao, M.G., Ramkumar, J., Ding, L., Tang, Y.T., Hafalia, A.J., Nguyen, D.B., Gandhi, A.R., Lu, Y., Yue, H., Burford, N., Bandman, O., Tribouley, C.M., Lal, P.G., Recipon, S.A., Lu, D.A., Borowsky, M.L., Thornton, M., Swarraker, A., Thangavelu, K., Khan, F.A. and Ison, C.H. Human kinases	Qy	841 GGCTGACCTGTGACTGTGCTGAGTGGGCTGATTCCTTATGAGATGATTTATGGGAGA 900
AUTHORS		Db	895 GGCTGACCTGTGACTGTGCTGAGTGGGCTGATTCCTTATGAGATGATTTATGGGAGA 954
TITLE	Patent: WO 0233099-A 43 25-APR-2002;	Qy	901 TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTATGATTTCCAGCGG 960
JOURNAL	Incyte Genomics, Inc. (US)	Db	955 TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTATGATTTCCAGCGG 1014
FEATURES	Location/Qualifiers	Qy	961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1020
source	1..6298	Db	1015 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1074
ORIGIN		Qy	1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC 1080
Query Match	92.0%; Score 5666.4; DB 6; Length 6298;	Db	1075 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC 1134
Best Local Similarity	95.5%; Pred. No. 0;	Qy	1081 TCTAAAATTGACTGGAAACACATTCGTAACCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCAG 1140
Matches 5961; Conservative	0; Mismatches 36; Indels 243; Gaps 3;	Db	1135 TCTAAAATTGACTGGAAACACATTCGTAACCTCTCCCTCCCTCCCTCCCTCCCTCCAG 1194
Qy	1 ATGTTGAAGTTCAATATATGAGCGCGGAATCTTTTGGATGCTGTGCTGCTGAACCCCAAT 60	Qy	1141 TCTGACCATGACACTCCAAATTTTGTATGAACACAGAGAAATTCGTGGGTTTCATCCTCT 1200
Db	55 ATGTTGAAGTTCAATATATGAGCGCGGAATCTTTTGGATGCTGTGCTGCTGAACCCCAAT 114	Db	1195 TCTGACCATGACACTCCAAATTTTGTATGAACACAGAGAAATTCGTGGGTTTCATCCTCT 1254
Qy	61 GCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGAAACACCCCTTTATGACTCAA 120	Qy	1201 CCCTGCGAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGGTTTCG 1260
Db	115 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGAAACACCCCTTTATGACTCAA 174	Db	1255 CCCTGCGAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGGTTTCG 1314
Qy	121 CACGAGATGTCCTCTTTCGCCAGAGGATATAGATGCGCTCTTGTCTCTTTGAA 180	Qy	1261 TACAGCAGGCTGCGGATTTCTGTGATGATCTGAGTCTGTTGTGCGGGTCTGGACTCC 1320
Db	175 CACGAGATGTCCTCTTTCGCCAGAGGATATAGATGCGCTCTTGTCTCTTTGAA 234	Db	1315 TACAGCAGGCTGCGGATTTCTGTGATGATCTGAGTCTGTTGTGCGGGTCTGGACTCC 1374
Qy	181 GAATGAGTCAGCTGCTCTGATGAAGTAAAGCAGTGAAGCAATTTTGTCCGGAAGTAT 240	Qy	1321 CCTCCCAAGCTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGAGCTACAGAC 1380
Db	235 GAATGAGTCAGCTGCTCTGATGAAGTAAAGCAGTGAAGCAATTTTGTCCGGAAGTAT 294	Db	1375 CCTCCCAAGCTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGAGCTACAGAC 1434
Qy	241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGAGCTTCCGAGTCAGA 300	Qy	1381 TCTCAGCAGAGTGTCAAGATGGAGCAAGATGACCCGTTTACATCCGAGAGTGTCA 1440
Db	295 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGAGCTTCCGAGTCAGA 354	Db	1435 TCTCAGCAGAGTGTCAAGATGGAGCAAGATGACCCGTTTACATCCGAGAGTGTCA 1494
Qy	301 AGTCTTGTAGTTTGTGTCATTTTGTGAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG 360	Qy	1441 GAGTGTGAGGCTGTGCTTAGTTCAGAGAGGAGTGGAGCTGAAGGCTCTGTGAGCTCAGAGA 1500
Db	355 AGTCTTGTAGTTTGTGTCATTTTGTGAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG 414	Db	1495 GAGTGTGAGGCTGTGCTTAGTTCAGAGAGGAGTGGAGCTGAAGGCTCTGTGAGCTCAGAGA 1554
Qy	361 GACATCTATGATGAAGTGAAGAGAGAGAGGCTTTATTCGCCAGGAGCAGGTTTCA 420	Qy	1501 TCCCTCTCTGGAGCAGGACCTTGTACCTACATACAGAAATGCGAGTAGCTTTAAAGCGAGT 1560
Db	415 GACATCTATGATGAAGTGAAGAGAGAGAGGCTTTATTCGCCAGGAGCAGGTTTCA 474	Db	1555 TCCCTCTCTGGAGCAGGACCTTGTACCTACATACAGAAATGCGAGTAGCTTTAAAGCGAGT 1614
Qy	421 TTTTGTGAGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480	Qy	1561 TTGAGCAAGCACCGATGGAGGTGTCCCGAGGAGATGACAAAGCAGCTGACGTTCTCCAT 1620
Db	475 TTTTGTGAGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 534	Db	1615 TTGAGCAAGCACCGATGGAGGTGTCCCGAGGAGATGACAAAGCAGCTGACGTTCTCCAT 1674
Qy	481 CAGTATGCCCTTTCAGGACAAATATCACTTTATCTGATGAGGATATCAGCTGGAGGG 540	Qy	1621 GATATCAGAGCAGAGCCCGAGAGCTCCAAAGAAATCAAGAGCAGGAGTAGCAGGCTCAA 1680
Db	535 CAGTATGCCCTTTCAGGACAAATATCACTTTATCTGATGAGGATATCAGCTGGAGGG 594	Db	1675 GATATCAGAGCAGAGCCCGAGAGCTCCAAAGAAATCAAGAGCAGGAGTAGCAGGCTCAA 1734
Qy	541 GACTTGTGTCATTTTGAATAGATATGAGGACAGTTCAGTGAAGAAACCTGTATACAGTTT 600		

Db	3893	-----AGGTTCTCTGCAGTAC	3909	Db	4930	CCCTTCAGTACACAGGTGCTGTTGGTGGGACCCAGGAAGGGCTCTACGCCCTGAATGTC	4989
Qy	3853	AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC	3912	Qy	4783	TTGAAAAATCCCTTAACCCATGTCCAGGAAATTTGGAGCAGTCTTCCAAATTTATATATTC	4842
Db	3910	AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC	3969	Db	4990	TTGAAAAATCCCTTAACCCATGTCCAGGAAATTTGGAGCAGTCTTCCAAATTTATATATTC	5049
Qy	3913	CTTTAGAAGACCCGATCAGAGCTCCGGTCCCGCCGCGGAGGAAGCTGCCCAACCGAAAGCA	3972	Qy	4843	AAGGACCTGAGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTTGGGAC	4902
Db	3970	CTTTAGAAGACCCGATCAGAGCTCCGGTCCCGCCGCGGAGGAAGCTGCCCAACCGAAAGCA	4029	Db	5050	AAGGACCTGAGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTTGGGAC	5109
Qy	3973	ACGGACACCCACACCATTCACGCCAGCACCGGAGGAGCAGAGATCGCATGTCGCC	4032	Qy	4903	GTGAAGAAAGTGAAGAACAGTCCCTGGGCCAGTCCACCTGCCTGCCAGCCCGACATCTCA	4962
Db	4030	ACGGACACCCACACCATTCACGCCAGCACCGGAGGAGCAGAGATCGCATGTCGCC	4089	Db	5110	GTGAAGAAAGTGAAGAACAGTCCCTGGGCCAGTCCACCTGCCTGCCAGCCCGACATCTCA	5169
Qy	4033	ATCGTCCGGTCCGACAGACACAGCCCACTGATGAGCTGCTGGCCCCCGCCATCCAGC	4092	Qy	4963	CCCAACATTTTGAAGTGTCAAGGGTCCCACTGTGTTGGGCGAGCAAGATTTGAGAAC	5022
Db	4090	ATCGTCCGGTCCGACAGACACAGCCCACTGATGAGCTGCTGGCCCCCGCCATCCAGC	4149	Db	5170	CCCAACATTTTGAAGTGTCAAGGGTCCCACTGTGTTGGGCGAGCAAGATTTGAGAAC	5229
Qy	4093	CGCAGAAAGAGTCTTCAATCTCAGAGGAATTTAGTCGCGCTTTAAGGAACGATGAC	4152	Qy	5023	GGGCTCTGCTCTGTGAGCATGCTCCAGCAATGCTCATTTCTCGCTACACAGAAAC	5082
Db	4150	CGCAGAAAGAGTCTTCAATCTCAGAGGAATTTAGTCGCGCTTTAAGGAACGATGAC	4209	Db	5230	GGGCTCTGCTCTGTGAGCATGCTCCAGCAATGCTCATTTCTCGCTACACAGAAAC	5289
Qy	4153	CACAATATTTCTCACCAGATTTCAACGTAGGACTGAACATCGAGGCCACAAAGTGTCTGTG	4212	Qy	5083	CTCAGCAAAATCTGCATTCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC	5142
Db	4210	CACAATATTTCTCACCAGATTTCAACGTAGGACTGAACATCGAGGCCACAAAGTGTCTGTG	4269	Db	5290	CTCAGCAAAATCTGCATTCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC	5349
Qy	4213	TGPTCTGGATACCGTGACATTTGGAACCCAGGATCCAAATGTCTCGAATGTCAGTGATG	4272	Qy	5143	TTCAACCAATTTACAGTATCCCTTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG	5202
Db	4270	TGPTCTGGATACCGTGACATTTGGAACCCAGGATCCAAATGTCTCGAATGTCAGTGATG	4329	Db	5350	TTCAACCAATTTACAGTATCCCTTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG	5409
Qy	4273	TGTCACCCCAAGTCTCCAGCTGCTTCCAGCACCTCGGCTTGCCTGCTCAATATGCC	4332	Qy	5203	TACACGCTCGAGGAATTTCTCGATAGAATGACCAATTCCTTTGGCACTGTGTGTGTC	5262
Db	4330	TGTCACCCCAAGTCTCCAGCTGCTTCCAGCACCTCGGCTTGCCTGCTCAATATGCC	4389	Db	5410	TACACGCTCGAGGAATTTCTCGATAGAATGACCAATTCCTTTGGCACTGTGTGTGTC	5469
Qy	4333	ACACACTTACCCAGGCTTCTGCGGTGACAAATGAACCTCCAGGTCTCCAGACCAAG	4392	Qy	5263	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCCAGGCGAGCAGAG	5322
Db	4390	ACACACTTACCCAGGCTTCTGCGGTGACAAATGAACCTCCAGGTCTCCAGACCAAG	4449	Db	5470	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCCAGGCGAGCAGAG	5529
Qy	4393	GAGCCAGACGAGCTTGACCTTGAAGGGTGAATGATGCTGCGGATCAAAAGTCTCATTTAT	4452	Qy	5323	GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTGTGGAATTTTACGGAAGACGTAGC	5382
Db	4450	GAGCCAGACGAGCTTGACCTTGAAGGGTGAATGATGCTGCGGATCAAAAGTCTCATTTAT	4509	Db	5530	GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTGTGGAATTTTACGGAAGACGTAGC	5589
Qy	4453	GGACAGCAAGCTGGACAGGAGTACATTTGCTGAGGAGTCAAAAGTCTCATTTAT	4512	Qy	5383	CGCAGACGATCTCAAGTGGAGTCTGTCTTACCTTTGGCTTTCCTACAGAGAACCTAT	5442
Db	4510	GGACAGCAAGCTGGACAGGAGTACATTTGCTGAGGAGTCAAAAGTCTCATTTAT	4569	Db	5590	CGCAGACGATCTCAAGTGGAGTCTGTCTTACCTTTGGCTTTCCTACAGAGAACCTAT	5649
Qy	4513	GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAAGATTTGAGCTGTGCTTCC	4572	Qy	5443	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTAAGGTCATTTGCTGC	5502
Db	4570	GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAAGATTTGAGCTGTGCTTCC	4629	Db	5650	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTAAGGTCATTTGCTGC	5709
Qy	4573	GACGGGATGTATCTATTCAATGTCGCTTGTGCTTCCGAACTCCCAATACAGCCAAA	4632	Qy	5503	GGGACCCCTCGCCGAGCGTACTTGGACATCCGAAACCCGCGCTACTTGGGCGCTGCCATT	5562
Db	4630	GACGGGATGTATCTATTCAATGTCGCTTGTGCTTCCGAACTCCCAATACAGCCAAA	4689	Db	5710	GGGACCCCTCGCCGAGCGTACTTGGACATCCGAAACCCGCGCTACTTGGGCGCTGCCATT	5769
Qy	4633	GCA-----	4635	Qy	5563	TCCTCAGAGCGATTTACTTGGGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC	5622
Db	4690	GCAGATGTCCCATACATAGATGGAATCTCACCCGACACACCTGCTGGCCCGG	4749	Db	5770	TCCTCAGAGCGATTTACTTGGGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC	5829
Qy	4636	-----	4635	Qy	5623	AAGGGAACCTCGTGAAGGAGTCCGGCACTGAAACACACCGGGGCGCTTCACTCCCGC	5682
Db	4750	AGAAACCTCTATTGCTAGCTCCAGCTTCCCTGACAAACAGGCTGGGTACCCGCTTCA	4809	Db	5830	AAGGGAACCTCGTGAAGGAGTCCGGCACTGAAACACACCGGGGCGCTTCACTCCCGC	5889
Qy	4636	-----GAAAAAGCAGAGCTGATGCTAAATCTG	4662	Qy	5883	AGCAGCCCAACAGCAGGCGCCACCCAGTACAAACAGCAGCATCACCAAGCGCTGGCC	5742
Db	4810	GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAAGCAGAGCTGATGCTAAATCTG	4869	Db	5890	AGCAGCCCAACAGCAGGCGCCACCCAGTACAAACAGCAGCATCACCAAGCGCTGGCC	5949
Qy	4663	CTTGGAAATCTCCCTGTGAAACTGGAAAGTGTAGCGCTTAGACATGAATGACGCTG	4722	Qy	5743	TCCAGCCAGCGCCCGCCAGAGCCCGCAGCCACCCGCGAGAGCCACACCCCAACCGC	5802
Db	4870	CTTGGAAATCTCCCTGTGAAACTGGAAAGTGTAGCGCTTAGACATGAATGACGCTG	4929	Db	5950	TCCAGCCAGCGCCCGCCAGAGCCCGCAGCCACCCGCGAGAGCCACACCCCAACCGC	6009
Qy	4723	CCCTTCAGTGACAGGTGGTGTGGTGGGACCGAGGAAGGGCTTTACGCCCTTGAATGTC	4782	Qy	5803	TACCGAGGCGGCGGACCGAGCTGCGCAGGACAGTCTCTTGGCGCCCTTGGAGCGA	5862
				Db	6010	TACCGAGGCGGCGGACCGAGCTGCGAGGGAAGTCTCTGCGCCGCCCTTGGAGCGA	6069

QY 5863 GAGAGTCCCGCGCGGATGCTCAGACGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5922
Db 6070 GAGAGTCCCGCGCGGATGCTCAGACGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6129
QY 5923 GAAGACAGCAGAGCGGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCACAGTG 5982
Db 6130 GAAGACAGCAGAGCGGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCACAGTG 6189
QY 5983 AACAGGAGAGGAGAGAGTCCCTCAAGTTTTCACGTTTAAACACTGTCACTATTAT 6042
Db 6190 AACAGGAGTGGGACCACTTTCAGTATAATCTCAGCCAGAAAAACCAACTCTCTCATCT 6249

RESULT 6
AX671105
LOCUS AX671105 6155 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO03004523.
ACCESSION AX671105
VERSION AX671105.1 GI:29329569
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS
TITLE Regulation of human citron rho/rac-interacting kinase
JOURNAL Patent: WO 03004523-A 1, 16-JAN-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1..6155
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 92.0%; Score 5666.2; DB 6; Length 6165;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 5947; Conservative 0; Mismatches 13; Indels 243; Gaps 3;

QY 1 ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAAACCCATT 60
Db 1 ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAAACCCATT 60
QY 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCTTTATGACTCAA 120
Db 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCTTTATGACTCAA 120
QY 121 CAGCAGATGTCCTCTTTCCGAGAGGATATAGATGCCCTCTTTGTTCTTTTGAA 180
Db 121 CAGCAGATGTCCTCTTTCCGAGAGGATATAGATGCCCTCTTTGTTCTTTTGAA 180
QY 181 GAATCAGTCAAGCTGCTCTGATGAAGATTAAAGCAGTGAACCTTTGTCGGGAAGTAT 240
Db 181 GAATCAGTCAAGCTGCTCTGATGAAGATTAAAGCAGTGAACCTTTGTCGGGAAGTAT 240
QY 241 TCCGACACATAGCTGAGTTACAGAGCTCAGCCTTCGGCAAGGACTTCGAAAGTACA 300
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCAGCCTTCGGCAAGGACTTCGAAAGTACA 300
QY 301 AGTCCTGAGTGTGCTGCTCCTGAGTGAAGTGAAGTGAAGAGAGAGAGAGAGAGAG 360
Db 301 AGTCCTGAGTGTGCTGCTCCTGAGTGAAGTGAAGTGAAGAGAGAGAGAGAGAGAG 360
QY 361 GACATCTATGCTATGAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 GACATCTATGCTATGAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TTTTGTGAG 480
Db 421 TTTTGTGAG 480

QY 481 CAGTATGCTTTCAGGACAAATATCAGCTTATCTGATGAGGAGATATCAGCTGGAGG 540
Db 481 CAGTATGCTTTCAGGACAAATATCAGCTTATCTGATGAGGAGATATCAGCTGGAGG 540
QY 541 GACTTGTGTCACCTTTTGAATAGATAGAGACAGTATGATGATGATGATGATGATGAT 600
Db 541 GACTTGTGTCACCTTTTGAATAGATAGAGACAGTATGATGATGATGATGATGATGAT 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGGTTCATCTGATGAGGATGATGATGAT 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGGTTCATCTGATGAGGATGATGATGAT 660
QY 661 GACATCAAGCTCAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
Db 661 GACATCAAGCTCAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
QY 721 GGAATCCCGGAAAATGAATTCAAAAGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 GGAATCCCGGAAAATGAATTCAAAAGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 CCAGATTACATGCTCTGAAAGTCTGACTGTGATGAAAGGGGATGAAAGGACACCTAC 840
Db 781 CCAGATTACATGCTCTGAAAGTCTGACTGTGATGAAAGGGGATGAAAGGACACCTAC 840
QY 841 GGCTGGAAGTGTGACTGTGCTGAGTGGGCTGATTCCTTATGAGATGATTTATGGAGA 900
Db 841 GGCTGGAAGTGTGACTGTGCTGAGTGGGCTGATTCCTTATGAGATGATTTATGGAGA 900
QY 901 TCCCTCTTCAGAGAGAGAGCTTCCAGAACCTTCAATACATGATGATTTCCAGCGG 960
Db 901 TCCCTCTTCAGAGAGAGAGCTTCCAGAACCTTCAATACATGATGATTTCCAGCGG 960
QY 961 TTTTGTGAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TTTTGTGAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 TTGTTGTGGGCGCAGAGAGAGAGCTGAAAGTCTTGAAGTCTTTGCTGCTGCTTCTTCT 1080
Db 1021 TTGTTGTGGGCGCAGAGAGAGAGCTGAAAGTCTTGAAGTCTTTGCTGCTGCTTCTTCT 1080
QY 1081 TCTAAATTTGACTGGAACCAACATTCGTAACCTTCCCTCCCTTCTGTTCCACCTCAAG 1140
Db 1081 TCTAAATTTGACTGGAACCAACATTCGTAACCTTCCCTCCCTTCTGTTCCACCTCAAG 1140
QY 1141 TCTGACGATGACACCTTCCAAATTTGATGAACCAAGAGAGAGATTCGTTGCTCTCT 1200
Db 1141 TCTGACGATGACACCTTCCAAATTTGATGAACCAAGAGAGAGATTCGTTGCTCTCT 1200
QY 1201 CCGTCCAGTGAAGCTCAGGCTTCTCGGTTGAAGAGTCCGCTTGTGGGTTTTCG 1260
Db 1201 CCGTCCAGTGAAGCTCAGGCTTCTCGGTTGAAGAGTCCGCTTGTGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 TACAGCAAGGCACTGGGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 CTTGCCAAGCTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGCTGATCAAGAC 1380
Db 1321 CTTGCCAAGCTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGCTGATCAAGAC 1380
QY 1381 TCTCAGGACAAAGTGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 TCTCAGGACAAAGTGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 GAGGTGAGGCTGCTGTTAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 GAGGTGAGGCTGCTGTTAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 TCCCTCTTCAG 1560
Db 1501 TCCCTCTTCAG 1560
QY 1561 TTTGAGCAAGCAGGATGGAGGTGTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620

1561 TTGGAGCAAGCGGATGAGGTGTCCTCCAGGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1620
1621 GATATCAGAGAGCAGAGCGGAGGCTCCAAAGAAATCAAAAGAGCAGGAGTACCAGGCTCAA 1680
1621 GATATCAGAGAGCAGAGCGGAGGCTCCAAAGAAATCAAAAGAGCAGGAGTACCAGGCTCAA 1680
1681 GTGGAAAGAAATGAGGTTGATGATGAATCACTGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740
1681 GTGGAAAGAAATGAGGTTGATGATGAATCACTGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740
1741 CGGAGTGATCTCTACCAATCTCAGCTGAGAGAGTCTCGGCTTTCGCTGCAAGAAATCAAG 1800
1741 CGGAGTGATCTCTACCAATCTCAGCTGAGAGAGTCTCGGCTTTCGCTGCAAGAAATCAAG 1800
1801 CGAAAGCGACAGAAATGTGAGCAATAAATCTGTTGAAGGCTTAAGATCAAGGGAAGCTGAA 1860
1801 CGAAAGCGACAGAAATGTGAGCAATAAATCTGTTGAAGGCTTAAGATCAAGGGAAGCTGAA 1860
1861 GTGGGAGAAATATCGGAAACTTGGAGAGATCAATGCTGACGACAGCTCAAAATTCAGGAG 1920
1861 GTGGGAGAAATATCGGAAACTTGGAGAGATCAATGCTGACGACAGCTCAAAATTCAGGAG 1920
1921 CTCGAAGAGAACTGGAGAAGGCT----- 1944
1921 CTCGAAGAGAACTGGAGAAGGCTGTAAAAGCCAGCAGCGAGGCCACCGAGCTGTGCGAG 1980
1945 -----GCAAGAGCGGAGCGGAGGAGCTGGAGAGCTGCGAGACCGGAG 1992
1981 AATATCCGCGCAAGAGCGGAGCGGAGGAGCTGGAGAGCTGCGAGACCGGAG 2040
1993 GATTTCTTCAAGGCTCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCCATTTCTCTG 2052
2041 GATTTCTTCAAGGCTCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCCATTTCTCTG 2100
2053 GAGAAACAAGTAAGAGCTAGAGACCATGGAGCGGTAGAGAAAACAGACTGAAGATGAC 2112
2101 GAGAAACAAGTAAGAGCTAGAGACCATGGAGCGGTAGAGAAAACAGACTGAAGATGAC 2160
2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGGAGCTCCAAGAG 2172
2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGGAGCTCCAAGAG 2220
2173 AAAACATCGGAGGCCCAAGTCTCAGCCAGACCTAGAAAGTCCACTGAAAACAGAAAGAG 2232
2221 AAAACATCGGAGGCCCAAGTCTCAGCCAGACCTAGAAAGTCCACTGAAAACAGAAAGAG 2280
2233 CAGCACTATCAGAGAAAGATTAAGTGTGGACATCAGATAAGAAAGACCTGGCTGAC 2292
2281 CAGCACTATCAGAGAAAGATTAAGTGTGGACATCAGATAAGAAAGACCTGGCTGAC 2340
2293 AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGGCCCATGAGAGGCGAAA 2352
2341 AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGGCCCATGAGAGGCGAAA 2400
2353 ATTCTCAGCGAAACAGAAAGGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2412
2401 ATTCTCAGCGAAACAGAAAGGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2460
2413 CAGAGGATTTGTGAACTGTCTGAAGCCAAATAACTTTGCAGCAAAATAGCAGTCTTTTACC 2472
2461 CAGAGGATTTGTGAACTGTCTGAAGCCAAATAACTTTGCAGCAAAATAGCAGTCTTTTACC 2520
2473 CAAAGGAAACATGAAGGCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
2521 CAAAGGAAACATGAAGGCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2592
2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2640
2593 AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTCTCGAACTGGAGACAGATTTG 2652

2641 AAGATCAGCCCAAGACCAACACAGTACAGATCGGCTGCTGGAACTGGAGACAAGATTG 2700
2653 CGGAGGTCAGTCTTAGAGCAAGAGAGCAGAAACTGGAGTCTAAGCGCCAGCTCACAGAG 2712
2701 CGGAGGTCAGTCTTAGAGCAAGAGAGCAGAAACTGGAGTCTAAGCGCCAGCTCACAGAG 2760
2713 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTCACAGCCCTGACAGCTGACCGGGCG 2772
2761 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTCACAGCCCTGACAGCTGACCGGGCG 2820
2773 GCCCTGGAGACCAAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGAGAAGCT 2832
2821 GCCCTGGAGACCAAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGAGAAGCT 2880
2833 GAAGAGGAGATCCAGGCACCTCAGGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2892
2881 GAAGAGGAGATCCAGGCACCTCAGGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940
2893 CTTTCGTAAACAGCTGTACTGTAAATCAAGACCTCGAGAGCAGCTAAACAGCTGACCGAG 2952
2941 CTTTCGTAAACAGCTGTACTGTAAATCAAGACCTCGAGAGCAGCTAAACAGCTGACCGAG 3000
2953 GACAACTGAACTCAACACCAAACTTCTACTTGTCCAAACAATCTCGATGAGGCTTCT 3012
3001 GACAACTGAACTCAACACCAAACTTCTACTTGTCCAAACAATCTCGATGAGGCTTCT 3060
3013 GGCCCAACCAAGAGATTGTACAACCTCGAAGTGAAGTGACCATCTCCGCCGGGAGATC 3072
3061 GGCCCAACCAAGAGATTGTACAACCTCGAAGTGAAGTGACCATCTCCGCCGGGAGATC 3120
3073 ACGAAACGAGAGATGAGCTTACCGCAGAAACCAAGATGAGAGCTCTGAAGACACAG 3132
3121 ACGAAACGAGAGATGAGCTTACCGCAGAAACCAAGATGAGAGCTCTGAAGACACAG 3180
3133 TGCACTCATGCTGGAGAAACAGGTCATGGATTTCGAGGCCCTAAACGATGAGCTGTAGAA 3192
3181 TGCACTCATGCTGGAGAAACAGGTCATGGATTTCGAGGCCCTAAACGATGAGCTGTAGAA 3240
3193 AAAGAGCGCAGTGGAGGCTTGGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTGAG 3252
3241 AAAGAGCGCAGTGGAGGCTTGGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTGAG 3300
3253 TGTGGGTTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3312
3301 TGTGGGTTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3360
3313 GATCAGCGGATCAACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGGAGCAAGGCT 3372
3361 GATCAGCGGATCAACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGGAGCAAGGCT 3420
3373 GAGATTCCTGCTTCGAGCAGGCTCTCAAGAGCAGAGCTGAGGCGGAGAGCCCTCTCT 3432
3421 GAGATTCCTGCTTCGAGCAGGCTCTCAAGAGCAGAGCTGAGGCGGAGAGCCCTCTCT 3480
3433 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3492
3481 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3540
3493 CAGCAGAGCTGGAGACTGAAACAGAGCTCAAAACAGAGGCTTTCGAGAGCAGGCAAA 3552
3541 CAGCAGAGCTGGAGACTGAAACAGAGCTCAAAACAGAGGCTTTCGAGAGCAGGCAAA 3600
3553 TTACAGCAGAGATGGACCTGACAGAAAATCAATTTTCCTGCTGACTCAAGGACTGCAA 3612
3601 TTACAGCAGAGATGGACCTGACAGAAAATCAATTTTCCTGCTGACTCAAGGACTGCAA 3660
3613 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATTCGAGTATCAGCTG 3672
3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATTCGAGTATCAGCTG 3720
3673 GAAGACATTCAGGTTCTCTATTTCTCATGAAAGTGAATGGAAGGCACTATTTCTCAA 3732
3721 GAAGACATTCAGGTTCTCTATTTCTCATGAAAGTGAATGGAAGGCACTATTTCTCAA 3780

QY 3733 CAAACCAAACTCATTTGATTTTCTGCAAGCAAAATGGACCAACCTGCTCTAAAGAAAAAG 3792
Db 3781 CAAACCAAACTCATTTGATTTTCTGCAAGCAAAATGGACCAACCTGCTCTAAAGAAAA - - 3838
QY 3793 GGTTTATTTAGTCGACGGAAGAGGACCCTGTTTACCCACACACAGGTTCTCTGCAAGTAC 3852
Db 3839 -----AGGTTCTCTGCAAGTAC 3855
QY 3853 AATGAGCTGAAGCTGGCCCTCGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAAAGCC 3912
Db 3856 AATGAGCTGAAGCTGGCCCTCGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAAAGCC 3915
QY 3913 CTTCAAGAACCCGCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCTGCCACCGCAAAAGCA 3972
Db 3916 CTTCAAGAACCCGCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCTGCCACCGCAAAAGCA 3975
QY 3973 AGGACCAACCCACCCATCCAGCCAGCCACCGGAGGCGACAGATGCCATGTCGCCG 4032
Db 3976 AGGACCAACCCACCCATCCAGCCAGCCACCGGAGGCGACAGATGCCATGTCGCCG 4035
QY 4033 ATCTGTCGGTCCGCGAGACAGCAGCCAGTGCATGAGCTCTGCGCCCGGCATCCAGC 4092
Db 4036 ATCTGTCGGTCCGCGAGACAGCAGCCAGTGCATGAGCTCTGCGCCCGGCATCCAGC 4095
QY 4093 CGCAGAAAGAGTCTTCAACTCCAGAGGAAATTTAGTCCGGCTTTAAGGAACGATGCAC 4152
Db 4096 CGCAGAAAGAGTCTTCAACTCCAGAGGAAATTTAGTCCGGCTTTAAGGAACGATGCAC 4155
QY 4153 CACAATATTTCTCACCGATTCACAGTTCAGACATGCAATGCGAGCCACAAAGTGTCTGTG 4212
Db 4156 CACAATATTTCTCACCGATTCACAGTTCAGACATGCAATGCGAGCCACAAAGTGTCTGTG 4215
QY 4213 TGTCTGGATACCGTGCATTTTGGAGCGCCAGGAGATCCAAATGTCTGAAATGTTCAGGTGATG 4272
Db 4216 TGTCTGGATACCGTGCATTTTGGAGCGCCAGGAGATCCAAATGTCTGAAATGTTCAGGTGATG 4275
QY 4273 TGTCAACCCAGTGTCCAGCTGCTTCCAGAGCCACTGGCGCTGCTGCTGAATATGCC 4332
Db 4276 TGTCAACCCAGTGTCCAGCTGCTTCCAGAGCCACTGGCGCTGCTGCTGAATATGCC 4335
QY 4333 ACACACTTCACCGAGGCTTTCTGCGGTGACAAAATGAATCCCCAGGTTCCAGACCAG 4392
Db 4336 ACACACTTCACCGAGGCTTTCTGCGGTGACAAAATGAATCCCCAGGTTCCAGACCAG 4395
QY 4393 GAGCCAGCAGAGCTTGCACCTGGAGAGGTGATGAAGGTGCCAGGATTAACAAACGA 4452
Db 4396 GAGCCAGCAGAGCTTGCACCTGGAGAGGTGATGAAGGTGCCAGGATTAACAAACGA 4455
QY 4453 GGACAGCAAGGCTGGACAGGAAGTACATTGTCTGGAGGATCAAAAGTCTCATTTAT 4512
Db 4456 GGACAGCAAGGCTGGACAGGAAGTACATTGTCTGGAGGATCAAAAGTCTCATTTAT 4515
QY 4513 GACATGAAGCCAGAGAACTGGACAGAGCCGGTGGAAAGATTTGAGCTGTGCCTTCCC 4572
Db 4516 GACATGAAGCCAGAGAACTGGACAGAGCCGGTGGAAAGATTTGAGCTGTGCCTTCCC 4575
QY 4573 GAGGGGATGTATCTATTATCATGTGCGCTTGGTGTCTCCGAACTCGCAATACAGCCAAA 4632
Db 4576 GAGGGGATGTATCTATTATCATGTGCGCTTGGTGTCTCCGAACTCGCAATACAGCCAAA 4635
QY 4633 GCA----- 4635
Db 4636 GCAGATGTCCATACATACATACTGAAGATGGAATCTCACCCGACACACCTGTGGCCCGGG 4695
QY 4636 ----- 4635
Db 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4755
QY 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662
Db 4756 GAATCAGTTGTCGACGGTGGAGAGTTTCTAGGANAAGACAGAGCTGATGCTAAACTG 4815

QY 4663 CTTGAAAACTCCCTGCTGAAAACTGGAAAGTGATGACCGTCTAGACATGAATGACCGCTG 4722
Db 4816 CTTGAAAACTCCCTGCTGAAAACTGGAAAGTGATGACCGTCTAGACATGAATGACCGCTG 4875
QY 4723 CCTTTCAGTCACAGAGTGGTGTGGTGGGACCGAGGAAGGCTCTACGCCCTGAATGTC 4782
Db 4876 CCTTTCAGTCACAGAGTGGTGTGGTGGGACCGAGGAAGGCTCTACGCCCTGAATGTC 4935
QY 4783 TTGAAAAAATCCCTTAACCCATGTGCCAGGAATTTGGAGCAGTCTTCCAAAATTTATATTATC 4842
Db 4936 TTGAAAAAATCCCTTAACCCATGTGCCAGGAATTTGGAGCAGTCTTCCAAAATTTATATTATC 4995
QY 4843 AAGACCTTGAGAAAGCTTACTCATGATAGCAGGAGAGAGCGGACACTGTCTCTGTGGAC 4902
Db 4996 AAGACCTTGAGAAAGCTTACTCATGATAGCAGGAGAGAGCGGACACTGTCTCTGTGGAC 5055
QY 4903 GTGAAGAAAGTGAACCAAGTCCCTGGGCCAGTCCACACCTGCTGCCAGCCCGACATCTCA 4962
Db 5056 GTGAAGAAAGTGAACCAAGTCCCTGGGCCAGTCCACACCTGCTGCCAGCCCGACATCTCA 5115
QY 4963 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCGAGCAAGATGAGAAC 5022
Db 5116 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCGAGCAAGATGAGAAC 5175
QY 5023 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTCTTCTCCGCTTACAACGAAAAC 5082
Db 5176 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTCTTCTCCGCTTACAACGAAAAC 5235
QY 5083 CTCAGCAAAATATGTCATTCGGAAGAGATAGAGACCTCAGAGCCCTGCAGTGTATCCAC 5142
Db 5236 CTCAGCAAAATATGTCATTCGGAAGAGATAGAGACCTCAGAGCCCTGCAGTGTATCCAC 5295
QY 5143 TTCAACCAATTACAGTATCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAACGAG 5202
Db 5296 TTCAACCAATTACAGTATCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAACGAG 5355
QY 5203 TACACGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGCACCTGTCTGTCTTGGC 5262
Db 5356 TACACGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGCACCTGTCTGTCTTGGC 5415
QY 5263 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAACAGCGAGGCGAGAGAG 5322
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAACAGCGAGGCGAGAGAG 5475
QY 5323 GAGTACTTGTCTGTCTTCCACGMAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5382
Db 5476 GAGTACTTGTCTGTCTTCCACGMAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535
QY 5383 CGCACAGACATCTCAAGTGGAGTCTTACCTTTGGCCTTTGCTTACAGAGAACCTTAT 5442
Db 5536 CGCACAGACATCTCAAGTGGAGTCTTACCTTTGGCCTTTGCTTACAGAGAACCTTAT 5595
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCACCGCTCCTCAGCA 5502
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCACCGCTCCTCAGCA 5655
QY 5503 GGGACCCCTGCGCGAGGCTACCTGGACATCCGAAACCCGGCTACCTGGGCGCTTGGCATT 5562
Db 5656 GGGACCCCTGCGCGAGGCTACCTGGACATCCGAAACCCGGCTACCTGGGCGCTTGGCATT 5715
QY 5663 TCCTCAGAGGGATTTACTTTGGCTCTCATACAGGATAAATTAAGGTTCATTTGTGTC 5622
Db 5716 TCCTCAGAGGGATTTACTTTGGCTCTCATACAGGATAAATTAAGGTTCATTTGTGTC 5775
QY 5623 AAGGGAACCTCTGCTGAAGAGTCCGGCACTGAACACACCCGGGCGCGCTCCACCTCCCGC 5682
Db 5776 AAGGGAACCTCTGCTGAAGAGTCCGGCACTGAACACACCCGGGCGCGCTCCACCTCCCGC 5835
QY 5683 AGCAGCCCAACAAAGCAGGCGCCACCCAGTAAACAGAGACATCAACAAAGCGGTGGCC 5742
Db 5836 AGCAGCCCAACAAAGCAGGCGCCACCCAGTAAACAGAGACATCAACAAAGCGGTGGCC 5895
QY 5743 TCCAGCCAGCGCCCGGAGGAGGCGCCAGCCACCCGCGAGAGCCCAAGCACACCCACCGC 5802

Db	5896		TCCAGCCACAGCGCGCCGCGAAGCGCCAGCCACCCGCGAGAGCGCAACGCCACCCACCGC	5895
Qy	5803		TACCCGACGGGCGGACCGAGCTCGCAGGACAAAGTCTCTCTGGCCGCCCCCTGGAGCGA	5862
Db	5956		TACCCGCGAGGGCGGACCGAGCTCGCAGGACAAAGTCTCTCTGGCCGCCCCCTGGAGCGA	6015
Qy	5863		GAGAAGTCCCCCGCGCGGATGCTCAGCACGCGGAGAGAGCGGTCCCCCGGAGAGCTGTTT	5922
Db	6016		GAGAAGTCCCCCGCGCGGATGCTCAGCACGCGGAGAGAGCGGTCCCCCGGAGAGCTGTTT	6075
Qy	5923		GAACACAGCAGCAGGGCGCGCTCCCTGCGGAGCCGTGAGGACCCCGCTGTCCCAGGTG	5982
Db	6076		GAACACAGCAGCAGGGCGCGCTCCCTGCGGAGCCGTGAGGACCCCGCTGTCCCAGGTG	6135
Qy	5983		AACAAGGGAAGAGGCGAGGTGC	6005
Db	6136		AACAAGGCTGGGACCAAGTCTTC	6158
RESULT 7				
AX574425				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 5944; Conservative				
0; Mismatches 16; Indels 243; Gaps 3;				
Qy	1	ATCTTGAGTTCAAATATGAGCGCGAATCTTCTGGATGCTGCTGCTGTAACCCATT	60	
Db	1	ATGTTGAAGTTCAAAATGAGCGCGAATCTTCTGGATGCTGCTGCTGTAACCCATT	60	
Qy	61	GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA	120	
Db	61	GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA	120	
Qy	121	CAGCAGATGCTCTCTTTCCGAGAGGGAATATTAGATGCCCTCTTGTCTCTTTGAA	180	
Db	121	CAGCAGATGCTCTCTTTCCGAGAGGGAATATTAGATGCCCTCTTGTCTCTTTGAA	180	
Qy	181	GAATGCAGTCAGCTGCTCTGTATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGAAGTAT	240	
Db	181	GAATGCAGTCAGCTGCTCTGTATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGAAGTAT	240	
Qy	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAGACTTCGAAGTCAGA	300	
Db	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAGACTTCGAAGTCAGA	300	
Qy	301	AGTCTTGTAGTTGTGGTCACTTTGCTGAAGTCAGGTGTTAAGAGAGAAGCAACCGGG	360	
Db	301	AGTCTTGTAGTTGTGGTCACTTTGCTGAAGTCAGGTGTTAAGAGAGAAGCAACCGGG	360	
Qy	361	GACATCTATGCTATGAAGTGAAGAAAGAGGCTTTATTGCGCCAGGACAGTTCATCA	420	

QY 1501 TCCCTCTGGACGAGACCTTCTGCTACCTACATCAGCAATGCAATGCTTAAAGCGAAGT 1560
Db 1501 TCCCTCTGGACGAGACCTTCTGCTACCTACATCAGCAATGCAATGCTTAAAGCGAAGT 1560
QY 1561 TTGGAGCAAGCAGGATGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
Db 1561 TTGGAGCAAGCAGGATGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
QY 1621 GATATCAGAGACAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680
Db 1621 GATATCAGAGACAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGGTTGATGATGATCAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAAAGA 1740
Db 1681 GTGGAAGAAATGAGGTTGATGATGATCAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAAAGA 1740
QY 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1800
Db 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1800
QY 1801 CGGAAGCCACAGAAATGTCAGCATAACTGTTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1860
Db 1801 CGGAAGCCACAGAAATGTCAGCATAACTGTTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1860
QY 1861 GTGGAGAAATATCGGAATCTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
Db 1861 GTGGAGAAATATCGGAATCTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
QY 1921 CTCCAGAGAAATCGAGAGAGCT----- 1944
Db 1921 CTCCAGAGAAATCGAGAGAGCTTAAAGCCAGCAGGAGCCAGGAGCTGTGCGAG 1980
QY 1945 -----GCAAGAGGAGGAGCCGAGAGGAGCTGGAAGCTGAGCAACCGAGAG 1992
Db 1981 AATATCCGCCAGGCANAGAGCGAGCCGAGAGGAGCTGAGAGCTGAGCAACCGAGAG 2040
QY 1993 GATTTCTTGAAGGCTATCAGAAAGAGCTGTGGAGAGCTGAGAGAGCCGCCATTTCTGTG 2052
Db 2041 GATTTCTTGAAGGCTATCAGAAAGAGCTGTGGAGAGCTGAGAGAGCCGCCATTTCTGTG 2100
QY 2053 GAGAACAGGTAAAGAGCTAGAGACCATGGAGCGTAGAGAAACAGACTGAAGGATGAC 2112
Db 2101 GAGAACAGGTAAAGAGCTAGAGACCATGGAGCGTAGAGAAACAGACTGAAGGATGAC 2160
QY 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTGAGGCTGCAAGAG 2172
Db 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTGAGGCTGCAAGAG 2220
QY 2173 AAACATCGGAGGCCCAAGCTCTCAGCCAGCACCCTAGAGAGTGCACCTGAAACAGAAAGAG 2232
Db 2221 AAACATCGGAGGCCCAAGCTCTCAGCCAGCACCCTAGAGAGTGCACCTGAAACAGAAAGAG 2280
QY 2233 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC 2292
Db 2281 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC 2340
QY 2293 AAGGAGCACTGAGAAACATGATGAGAGACAGAGGAGAGGCCCATGAGAGGGCAAA 2352
Db 2341 AAGGAGCACTGAGAAACATGATGAGAGACAGAGGAGAGGCCCATGAGAGGGCAAA 2400
QY 2353 ATTCTCAGGCAACAGAGGCGATGATCAATGCTATGATTTCCAAAGATCAGATCCCTGGAA 2412
Db 2401 ATTCTCAGGCAACAGAGGCGATGATCAATGCTATGATTTCCAAAGATCAGATCCCTGGAA 2460
QY 2413 CAGAGGATTTGGAATGCTGAGGCCAATPAAATCTTGCAGCAAAATAGCAGTCTTTTACC 2472
Db 2461 CAGAGGATTTGGAATGCTGAGGCCAATPAAATCTTGCAGCAAAATAGCAGTCTTTTACC 2520
QY 2473 CAAAGGACATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAAACAGAAATTTTAC 2532
Db 2521 CAAAGGACATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAAACAGAAATTTTAC 2580

QY 2533 CTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAACCGAAAACTGGAGGAGCAGCTGGAG 2592
Db 2581 CTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAACCGAAAACTGGAGGAGCAGCTGGAG 2640
QY 2593 AAGATCAGCCACCAAGCCACAGTGCAGAAATCGGCTCTGGAATCTGGAGAGCAAGATTG 2652
Db 2641 AAGATCAGCCACCAAGCCACAGTGCAGAAATCGGCTCTGGAATCTGGAGAGCAAGATTG 2700
QY 2653 CGGAGGTCTAGTCTTAGAGCACGAGGAGCAAGAACTTGAGCTCAAGCCAGCTCAGCAGAG 2712
Db 2701 CGGAGGTCTAGTCTTAGAGCACGAGGAGCAAGAACTTGAGCTCAAGCCAGCTCAGCAGAG 2760
QY 2713 CTAAGCTCTCCCTGAGGAGCGGAGTCACTGATTGACAGCCCTGAGGCTGACCGGGCG 2772
Db 2761 CTAAGCTCTCCCTGAGGAGCGGAGTCACTGATTGACAGCCCTGAGGCTGACCGGGCG 2820
QY 2773 GCCCTGAGAGCCAGCTTCGCCAGGGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2832
Db 2821 GCCCTGAGAGCCAGCTTCGCCAGGGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2880
QY 2833 GAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCGCAAAATTTGATGCT 2892
Db 2881 GAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCGCAAAATTTGATGCT 2940
QY 2893 CTTCTTAACAGCTGTACTGTAAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 2952
Db 2941 CTTCTTAACAGCTGTACTGTAAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 3000
QY 2953 GACAACGCTGAATCTCAACAAACCAAACTTCTTACTTGTCCAAACAACTCGATGAGGCTTCT 3012
Db 3001 GACAACGCTGAATCTCAACAAACCAAACTTCTTACTTGTCCAAACAACTCGATGAGGCTTCT 3060
QY 3013 GGCCTCAACGAGATTTGAACTGCGAAGTGAAGTGAACCACTCTCCGCCGGAGATC 3072
Db 3061 GGCCTCAACGAGATTTGAACTGCGAAGTGAAGTGAACCACTCTCCGCCGGAGATC 3120
QY 3073 ACGGAAACGAGAGATGACAGCTTACCAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3132
Db 3121 ACGGAAACGAGAGATGACAGCTTACCAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3180
QY 3133 TGCAACCATGCTGGAGGAAACAGGTCTATGATTTGGAGCCCTTAAACGATGAGCTCTAGAA 3192
Db 3181 TGCAACCATGCTGGAGGAAACAGGTCTATGATTTGGAGCCCTTAAACGATGAGCTCTAGAA 3240
QY 3193 AAAAGGCGGAGTGGGAGGCTGGAGGAGGCTCTGGGTGATGAGAAATCCAGCTTTGAG 3252
Db 3241 AAAAGGCGGAGTGGGAGGCTGGAGGAGGCTCTGGGTGATGAGAAATCCAGCTTTGAG 3300
QY 3253 TGTCCGGTTTCAGAGCTGACAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGAGCC 3312
Db 3301 TGTCCGGTTTCAGAGCTGACAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGAGCC 3360
QY 3313 GATCAGCGGATCAACCGAGTCTCGCCAGGTTGGAGCTGGAGTGAAGGAGCACAAGGCT 3372
Db 3361 GATCAGCGGATCAACCGAGTCTCGCCAGGTTGGAGCTGGAGTGAAGGAGCACAAGGCT 3420
QY 3373 GAGATTTCTGCTCTGACAGGCTCTCAAGAGCAGAAAGCTGAAAGGCGGAGAGCTCTCT 3432
Db 3421 GAGATTTCTGCTCTGACAGGCTCTCAAGAGCAGAAAGCTGAAAGGCGGAGAGCTCTCT 3480
QY 3433 GACAAGCTCAATGACCTTGAGAAAGAAATGCTATGTTGAAATGAATGCCCGAAGCTTA 3492
Db 3481 GACAAGCTCAATGACCTTGAGAAAGAAATGCTATGTTGAAATGAATGCCCGAAGCTTA 3540
QY 3493 CAGCAGAGCTGGAGCTGAAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAGCAAGCCAAA 3552
Db 3541 CAGCAGAGCTGGAGCTGAAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAGCAAGCCAAA 3600
QY 3553 TTACAGCAGCAGATGGAGCTTCAGAAAAATCACAATTTTCCGCTCTGACTCAAGGCTGCAA 3612
Db 3601 TTACAGCAGCAGATGGAGCTTCAGAAAAATCACAATTTTCCGCTCTGACTCAAGGCTGCAA 3660
QY 3613 GAAGCTCTAGATCGGGCTGATCTTACTGAGACAGAAAGAGTACTTGGAGTATCAGCTG 3672


```
QY 5683 AGCAGCCCCAACACGAGGCCACCCACGTAACAGAGACATCACCAAGCGGTGGCC 5742
DB 5836 AGCAGCCCCAACACGAGGCCACCCACGTAACAGAGACATCACCAAGCGGTGGCC 5895
QY 5743 TCCAGCCAGCCGCCCGAAGGCCCCAGCCACCCGCGAGAGCCAAAGCACACCCACCGC 5802
DB 5896 TCCAGCCAGCCGCCCGAAGGCCCCAGCCACCCGCGAGAGCCAAAGCACACCCACCGC 5955
QY 5803 TACCAGCGAGCGGCGACCGAGCTGGCGAGGGAACAAGTCTCTGGCGGCCCTCTGAGCGGA 5862
DB 5956 TACCAGCGAGCGGCGACCGAGCTGGCGAGGGAACAAGTCTCTGGCGGCCCTCTGAGCGGA 6015
QY 5863 GAGAGTCTCCCGCGGAGTGTCTAGCAGCGGAGAGCGGTCCCGGAGGCTGTT 5922
DB 6016 GAGAGTCTCCCGCGGAGTGTCTAGCAGCGGAGAGCGGTCCCGGAGGCTGTT 6075
QY 5923 GAAGCAGCAGCAGGCGCGGTGCTGGCGAGCGGTGAGGACCCCGCTGTCCAGGTG 5982
DB 6076 GAAGCAGCAGCAGGCGCGGTGCTGGCGAGCGGTGAGGACCCCGCTGTCCAGGTG 6135
QY 5983 AACAAAGGGAAGGCGCAGAGTGC 6005
DB 6136 AACAAAGTCTGGGACCAAGTCTTC 6158

RESULT 8
AXI166510
LOCUS AXI166510 6159 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1 from Patent WO0138503.
ACCESSION AXI166510
VERSION AXI166510.1 GI:14546855
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 1 31-MAY-2001;
Sugen, Inc. (US)
FEATURES
source
1..6159
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.7%; Score 5650.4; DB 6; Length 6159;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 5943; Conservative 0; Mismatches 11; Indels 246; Gaps 4;
QY 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTCTGTAACCCATT 60
DB 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTCTGTAACCCATT 60
QY 61 GCCAGCGCGCCCTCCAGGCTGAATCTGCTCTCCAGGGGAACACCCCTTTATGACTCAA 120
DB 61 GCCAGCGCGCCCTCCAGGCTGAATCTGTTCTTCAGGGGAACACCCCTTTATGACTCAA 120
QY 121 CAGCAGATGTCTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTTTTGAA 180
DB 121 CAGCAGATGTCTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTTTTGAA 180
QY 181 GAATGAGTACGCTGTCTGTATGAGATTAAAGCAGTGAACCTTTGTCGGGA--G 237
DB 181 GAATGAGTACGCTGTCTGTATGAGATTAAAGCAGTGAACCTTTGTCGGGAAGTG 240
QY 238 TATTCCGACACCAATAGTGTGATGAGTACAGGAGTCCAGGCTTCGGCAAGGACTTCGAAGTC 297
DB 241 TATTCCGACACCAATAGTGTGATGAGTACAGGAGTCCAGGCTTCGGCAAGGACTTCGAAGTC 300
```

```
QY 298 AGAAGTCTTGTAGTGTGCTGCTCACTTTGCTGAAGTGCAGGTGCTAAGAGAGAAGCAACC 357
DB 301 AGAAGTCTTGTAGTGTGCTGCTCACTTTGCTGAAGTGCAGGTGCTAAGAGAGAAGCAACC 360
QY 358 GGGGACATCTATGCTATGAAAGTGATGAGAAGAAGGCTTTATTGGCCCAAGGACAGGTT 417
DB 361 GGGGACATCTATGCTATGAAAGTGATGAGAAGAAGGCTTTATTGGCCCAAGGACAGGTT 420
QY 418 TCATTTTGGAGAGAGCGGAACATATTAATCTCGAAGCAACAGCCCGTGGATCCCCCAA 477
DB 421 TCATTTTGGAGAGAGCGGAACATATTAATCTCGAAGCAACAGCCCGTGGATCCCCCAA 480
QY 478 TTACAGTATCCCTTTTCAGGACAAAATACCTTTATCTGATGGAGGAATAATCAGCCTTGA 537
DB 481 TTACAGTATCCCTTTTCAGGACAAAATACCTTTATCTGATGGAGGAATAATCAGCCTTGA 540
QY 538 GGGGACCTTGCTGTCACTTTTGAATAGATATGAGGACCAAGTATGATAAAACCTTGATACAG 597
DB 541 GGGGACCTTGCTGTCACTTTTGAATAGATATGAGGACCAAGTATGATAAAACCTTGATACAG 600
QY 598 TTTTACCTAGCTGAGCTGATTTTGGCTGTTACAGCGTTTCATCTGATGGATAGCTGCAT 657
DB 601 TTTTACCTAGCTGAGCTGATTTTGGCTGTTACAGCGTTTCATCTGATGGATAGCTGCAT 660
QY 658 CGAGACATCAAGCCTGAGAAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGAT 717
DB 661 CGAGACATCAAGCCTGAGAAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGAT 720
QY 718 TTTGGATCTGCGGGGAAAATGAATTCAAACAAGATGGTGAATGCCAAACCTCCGATTTGGG 777
DB 721 TTTGGATCTGCGGGGAAAATGAATTCAAACAAGATGGTGAATGCCAAACCTCCGATTTGGG 780
QY 778 ACCCGAGATTACATGGCTCTCTGAAGTGTCTGATGATGAACGGGGATGAAAAAGGCACC 837
DB 781 ACCCGAGATTACATGGCTCTCTGAAGTGTCTGATGATGAACGGGGATGAAAAAGGCACC 840
QY 838 TACGGCTTGGACTGTGACTGCTGAGTGGCGGTGATTCGCTATAGATGATTTATGGG 897
DB 841 TACGGCTTGGACTGTGACTGCTGAGTGGCGGTGATTCGCTATAGATGATTTATGGG 900
QY 898 AGATCCCCCTTCGAGAGGGAACCTTCGCAGAACCTTCAATTAACATTTATGAATTTCCAG 957
DB 901 AGATCCCCCTTCGAGAGGGAACCTTCGCAGAACCTTCAATTAACATTTATGAATTTCCAG 960
QY 958 CGGTTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGTCTTCTGATCTGATTCAT 1017
DB 961 CGGTTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGTCTTCTGATCTGATTCAT 1020
QY 1018 AGCTTGTGTCGGCCAGAAAGAGAGACTGAAGTTGAAAGTGTCTTCTGCTCCATCCTTC 1077
DB 1021 AGCTTGTGTCGGCCAGAAAGAGAGACTGAAGTTGAAAGTGTCTTCTGCTCCATCCTTC 1080
QY 1078 TTCTCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCCCCCTTCGTTCCCAACCTTC 1137
DB 1081 TTCTCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCCCCCTTCGTTCCCAACCTTC 1140
QY 1138 AAGTCTGACGATGACACTCCCAATTTTGAATGAACACAGAGAAGATTCGTGGGTTTCATCC 1197
DB 1141 AAGTCTGACGATGACACTCCCAATTTTGAATGAACACAGAGAAGATTCGTGGGTTTCATCC 1200
QY 1198 TCTCCGTCAGCTGACGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTT 1257
DB 1201 TCTCCGTCAGCTGACGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTT 1260
QY 1258 TCGTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTGTGTGGGCTGTGAC 1317
DB 1261 TCGTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTGTGTGGGCTGTGAC 1320
QY 1318 TCCCTTGCAGAGACTAGCTCCATGGAAGAAATTTCTCATCAAAAGCAAGAGACTACAA 1377
DB 1321 TCCCTTGCAGAGACTAGCTCCATGGAAGAAATTTCTCATCAAAAGCAAGAGACTACAA 1380
```

QY 1378 GACTCTCAGGACAAAGTGTCAACAAGTGGAGCAGGAAATGACCCGGTTACATCGAGAGTG 1437
Db 1381 GACTCTCAGGACAAAGTGTCAACAAGTGGAGCAGGAAATGACCCGGTTACATCGAGAGTG 1440
QY 1438 TCAGAGGTGGAGGCTGTGCTTAGTTCAGAGAGAGGTGGAGTGAAGGCTCTGAGACTCAG 1497
Db 1441 TCAGAGGTGGAGGCTGTGCTTAGTTCAGAGAGAGGTGGAGTGAAGGCTCTGAGACTCAG 1500
QY 1498 AGATCCCTCTGGAGCAGGACCTTGTCTACATACACAGAAATGACAGTGTAAAGCGA 1557
Db 1501 AGATCCCTCTGGAGCAGGACCTTGTCTACATACACAGAAATGACAGTGTAAAGCGA 1560
QY 1558 AGTTTGGAGCAAGCAGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAAGCTTCTC 1617
Db 1561 AGTTTGGAGCAAGCAGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAAGCTTCTC 1620
QY 1618 CATGATATCAGAGCAGAGCCGAGCTCCAGAGAAATCAAGAGCAGAGTACCAAGGCT 1677
Db 1621 CATGATATCAGAGCAGAGCCGAGCTCCAGAGAAATCAAGAGCAGAGTACCAAGGCT 1680
QY 1678 CAAGTGAAGAAATGAGTTGATGATCAATCACTGTGGAAGAGGATCTTGTCTCAGCAAGA 1737
Db 1681 CAAGTGAAGAAATGAGTTGATGATCAATCACTGTGGAAGAGGATCTTGTCTCAGCAAGA 1740
QY 1738 AGACGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAGAAATTC 1797
Db 1741 AGACGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAGAAATTC 1800
QY 1798 AAGCGGAAACGACAGAAATGTCAAGCAATACTTGAAGGCTAAGGATCAAGGGAAGCCT 1857
Db 1801 AAGCGGAAACGACAGAAATGTCAAGCAATACTTGAAGGCTAAGGATCAAGGGAAGCCT 1860
QY 1858 GAAGTGGAGAAATATGCGAAATCTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1917
Db 1861 GAAGTGGAGAAATATGCGAAATCTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1920
QY 1918 GAGCTCCAGAGAAATCTGGAGAGGCT----- 1944
Db 1921 GAGCTCCAGAGAAATCTGGAGAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGTG 1980
QY 1945 -----GCAAGAGGACGAGCGGAGGAGCTGGAGAGGCTGGAGAACTGCAGAAACCGA 1989
Db 1981 CAGAAATATCCGACGCAAGAGAGCGAGCGAGGAGCTGGAGAGCTGCAGAAACCGA 2040
QY 1990 GAGGATCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGAGCTGAGGAAGCCGCCATTC 2049
Db 2041 GAGGATCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGAGCTGAGGAAGCCGCCATTC 2100
QY 2050 CTGGAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAAGACTGAAGAT 2109
Db 2101 CTGGAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAAGACTGAAGAT 2160
QY 2110 GACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAA 2169
Db 2161 GACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAA 2220
QY 2170 GAGAAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCCTGAAAACAGAAA 2229
Db 2221 GAGAAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCCTGAAAACAGAAA 2280
QY 2230 GAGCAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCT 2289
Db 2281 GAGCAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCT 2340
QY 2290 GACAGGAGACACTGGAGAACTGATGACAGACACGAGGAGGAGGCCATCAGAGAGGC 2349
Db 2341 GACAGGAGACACTGGAGAACTGATGACAGACACGAGGAGGAGGCCATCAGAGAGGC 2400
QY 2350 AAAATTTCTCAGGCAACAGAGGCGGATGATCAATGCTATGGATTCCAAGATCAGATCCCTG 2409
Db 2401 AAAATTTCTCAGGCAACAGAGGCGGATGATCAATGCTATGGATTCCAAGATCAGATCCCTG 2460
QY 2410 GAACAGAGATTTGTGAACTGTCTGAAGCCAAATAACTTGCAGCMAATAGCAGTCTTTT 2469

Db 2461 GAACAGAGATTTGTGAACTGTCTGAAGCCAAATAACTTTGAGCMAATAGCAGTCTTTT 2520
QY 2470 ACCCAAAAGAAACATGAAGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTT 2529
Db 2521 ACCCAAAAGAAACATGAAGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTT 2580
QY 2530 TACCTGGAGACACAGGCTGGAGCTTGGAGCCCAAGAACCCGAAACTCGAGAGCAGCTG 2589
Db 2581 TACCTGGAGACACAGGCTGGAGCTTGGAGCCCAAGAACCCGAAACTCGAGAGCAGCTG 2640
QY 2590 GAGAAAGATCAGCCACCAAGACACAGTGCAGAGAAATCGCTGCTGGAACCTGGAGACAGA 2649
Db 2641 GAGAAAGATCAGCCACCAAGACACAGTGCAGAGAAATCGCTGCTGGAACCTGGAGACAGA 2700
QY 2650 TTGCGGAGGTCAGTCTAGAGACAGAGAGCAGAAACTGGAGCTCAAGCGCCAGCTCA 2709
Db 2701 TTGCGGAGGTCAGTCTAGAGACAGAGAGCAGAAACTGGAGCTCAAGCGCCAGCTCA 2760
QY 2710 GAGCTTACAGCTCTCCCTCAGAGCGCGAGTCAAGTTGACAGCCCTGCGAGCTGCGG 2769
Db 2761 GAGCTTACAGCTCTCCCTCAGAGCGCGAGTCAAGTTGACAGCCCTGCGAGCTGCGG 2820
QY 2770 GCGGCCCTGGAGAGCAGCTTTCGCGAGCGAAGACAGAGCTGGAGAGACACAGCAGAA 2829
Db 2821 GCGGCCCTGGAGAGCAGCTTTCGCGAGCGAAGACAGAGCTGGAGAGACACAGCAGAA 2880
QY 2830 GCTGAAGAGGAGATTCAGGCACTCAGCGCATAGAGATGAAATCCAGCGCAATTTGAT 2889
Db 2881 GCTGAAGAGGAGATTCAGGCACTCAGCGCATAGAGATGAAATCCAGCGCAATTTGAT 2940
QY 2890 GCTCTTCGTAAACAGCTGTACTGTAAATCACAGACTGGAGGAGCAGCTTAAACAGCTGACC 2949
Db 2941 GCTCTTCGTAAACAGCTGTACTGTAAATCACAGACTGGAGGAGCAGCTTAAACAGCTGACC 3000
QY 2950 GAGGACAACTGTAATCAACACCAAACTTCTACTTGTCCAAAACACTTCGATGAGGCT 3009
Db 3001 GAGGACAACTGTAATCAACACCAAACTTCTACTTGTCCAAAACACTTCGATGAGGCT 3060
QY 3010 TCTGCGCCCAACGACGAGATTTGTACAACTCGAAGTGAAGTGAAGTCTCCCGCGGAG 3069
Db 3061 TCTGCGCCCAACGACGAGATTTGTACAACTCGAAGTGAAGTGAAGTCTCCCGCGGAG 3120
QY 3070 ATCAGGAAACAGAGATGCGAGCTTACCGCCAGAGCAACAGATGGAGGCTCTGAAGACC 3129
Db 3121 ATCAGGAAACAGAGATGCGAGCTTACCGCCAGAGCAACAGATGGAGGCTCTGAAGACC 3180
QY 3130 ACGTGCACCATGCTGGAGGAAACAGTCTCATGATTTGGAGGCCCTTAAACGATGAGCTGA 3189
Db 3181 ACGTGCACCATGCTGGAGGAAACAGTCTCATGATTTGGAGGCCCTTAAACGATGAGCTGA 3240
QY 3190 GAAAAAGAGCGGAGTGGAGGCTCGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTT 3249
Db 3241 GAAAAAGAGCGGAGTGGAGGCTCGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTT 3300
QY 3250 GAGTGTGCGGTTCCAGAGCTGCGAGAGTGTGCAACCGAGAGAAACAGAGCAGGCGGAGA 3309
Db 3301 GAGTGTGCGGTTCCAGAGCTGCGAGAGTGTGCAACCGAGAGAAACAGAGCAGGCGGAGA 3360
QY 3310 GCCGATCAGCGGATCACCGAGTCTCCCGAGGCTGGAGGCTGGAGGAGCAGCAAG 3369
Db 3361 GCCGATCAGCGGATCACCGAGTCTCCCGAGGCTGGAGGCTGGAGGAGCAGCAAG 3420
QY 3370 GCTGAGATTTCTCCCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGGCTC 3429
Db 3421 GCTGAGATTTCTCCCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGGCTC 3480
QY 3430 TCTGACAAAGCTCAATGACCTGGAGAGAGAGCTATGCTTATGAAATGAATGCCGAGGC 3489
Db 3481 TCTGACAAAGCTCAATGACCTGGAGAGAGAGCTATGCTTATGAAATGAATGCCGAGGC 3540
QY 3490 TTACAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGGCC 3549

Db 3541 TTACAGCAGAGCTGGAGACTGAACGAGAGCTCAACACAGAGGCTTCTGGAAGCAGCACC 3600
QY 3550 AATATTACAGCAGAGATGAGACCTGCGAGAAAAATCAATTTTCGGTCTGACTCAAGGACTG 3609
Db 3601 AATATTACAGCAGAGATGAGACCTGCGAGAAAAATCAATTTTCGGTCTGACTCAAGGACTG 3660
QY 3610 CAAGAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGACATTGAGATACAG 3669
Db 3661 CAAGAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGACATTGAGATACAG 3720
QY 3670 CTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAAGGTGAAAAATGGAAGGCACTATTCT 3729
Db 3721 CTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAAGGTGAAAAATGGAAGGCACTATTCT 3780
QY 3730 CAACAAACCAACTCATTTGATTTCTGCAAGCCAAATGGACCACTCTCTCTCTCTCTCTCT 3789
Db 3781 CAACAAACCAACTCATTTGATTTCTGCAAGCCAAATGGACCACTCTCTCTCTCTCTCTCT 3840
QY 3790 AAGGGTTTATTAGTCGAGAAAGAGGACCTGTCTTTACCCACACAGATTCTCTCTGAG 3849
Db 3841 A-----AGGTCTCTCTGAG 3855
QY 3850 TACAAATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGCTCGCTGTGCGAGACTAGAGAA 3909
Db 3856 TACAATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGCTCGCTGTGCGAGACTAGAGAA 3915
QY 3910 GCCCTTCAGAAAGACCCGCTCGAGTCCGGTCCGCCCGGAGAGAAAGCTGCCACCGCAAA 3969
Db 3916 GCCCTTCAGAAAGACCCGCTCGAGTCCGGTCCGCCCGGAGAGAAAGCTGCCACCGCAAA 3975
QY 3970 GNAACGAGCAACCAACCCATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4029
Db 3976 GCAACGAGCAACCAACCCATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4035
QY 4030 GCCATCGTGGGTGCGCAGAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 4089
Db 4036 GCCATCGTGGGTGCGCAGAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 4095
QY 4090 AGCCGAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACCGCATG 4149
Db 4096 AGCCGAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACCGCATG 4155
QY 4150 CACCACAAATTCCTCACCGATTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4209
Db 4156 CACCACAAATTCCTCACCGATTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4215
QY 4210 GTGTGTCTGGATACCGTGCACATTTGGACGCCAGGATCCAAATGTCTCGAATGTGAGGTG 4269
Db 4216 GTGTGTCTGGATACCGTGCACATTTGGACGCCAGGATCCAAATGTCTCGAATGTGAGGTG 4275
QY 4270 ATGTGTACCCCAAGTGTCTCCAGTGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4329
Db 4276 ATGTGTACCCCAAGTGTCTCCAGTGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 4335
QY 4330 GCCACACACTTACCGAGGCTTCTCCGCTGACAAATGAATCCCGAGGTCTCCAGACC 4389
Db 4336 GCCACACACTTACCGAGGCTTCTCCGCTGACAAATGAATCCCGAGGTCTCCAGACC 4395
QY 4390 AAGGAGCCAGCAGCAGTGTGACCTGGAAGGAGTGAATGAAGTGTGAGTGTGAGTGTGAGT 4449
Db 4396 AAGGAGCCAGCAGCAGTGTGACCTGGAAGGAGTGAATGAAGTGTGAGTGTGAGTGTGAGT 4455
QY 4450 CGAGGACAGAGGCTGGGACAGGAGTACATGTCTCGAGGATCAAAAGTCTCTATT 4509
Db 4456 CGAGGACAGAGGCTGGGACAGGAGTACATGTCTCGAGGATCAAAAGTCTCTATT 4515
QY 4510 TATGCAATGAAGCCAGAGAGTGGACAGAGGCGGTGGAGAAATTTGAGTGTGCTCT 4569
Db 4516 TATGCAATGAAGCCAGAGAGTGGACAGAGGCGGTGGAGAAATTTGAGTGTGCTCT 4575
QY 4570 CCGCAGGGGATGTATCTATTATGTGTGCGGTGCTCTCGAACTCCGAAATTCAGCC 4629
Db 4576 CCGCAGGGGATGTATCTATTATGTGTGCGGTGCTCTCGAACTCCGAAATTCAGCC 4635

QY 4630 AAAGCAGA----- 4637
Db 4636 AAAGCAGATGTCCCATACATACTGAAGATGAATCTCACCGCAGACACACCTGTGTGGCC 4695
QY 4638 ----- 4637
Db 4696 GGGAGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCC 4755
QY 4638 -----AAAGCAGAAGCTGATGCTAAA 4659
Db 4756 TTAGATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAAGCAGAGTGTGATGCTAAA 4815
QY 4660 CTGCTTGGAAACTCCCTGCTGAAACTGAAAGTGAATGATCAGCGTCTAGACATGAATGCAAG 4719
Db 4816 CTGCTTGGAAACTCCCTGCTGAAACTGAAAGTGAATGATCAGCGTCTAGACATGAATGCAAG 4875
QY 4720 CTGCCCCCTAGTGACACAGGTGTGTTGGTGGGACCCAGGAAGGGCTCTACGCCCTGAAT 4779
Db 4876 CTGCCCCCTAGTGACACAGGTGTGTTGGTGGGACCCAGGAAGGGCTCTACGCCCTGAAT 4935
QY 4780 GTCTTGAATAACTCCCTAAACCATGTCCAGGAAATGGAGCAGTCTTCCAAATTTATATT 4839
Db 4936 GTCTTGAATAACTCCCTAAACCATGTCCAGGAAATGGAGCAGTCTTCCAAATTTATATT 4995
QY 4840 ATCAAGGACCTGGAGAAAGTACTCATGATAGCAGAGAGAGCGGCACTGTGTCTGTG 4899
Db 4996 ATCAAGGACCTGGAGAAAGTACTCATGATAGCAGAGAGAGCGGCACTGTGTCTGTG 5055
QY 4900 GACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCTGCCAGCCGACATC 4959
Db 5056 GACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCTGCCAGCCGACATC 5115
QY 4960 TCACCCCAACATTTTGAAGCTGTCAAGGGTGCACATTTGTTGGGCGAGGCAAGATTGAG 5019
Db 5116 TCACCCCAACATTTTGAAGCTGTCAAGGGTGCACATTTGTTGGGCGAGGCAAGATTGAG 5175
QY 5020 AACGGGTCTGTGATCTGTGAGCCATGCCAGAGAAAGTGTGTCTCTCCGCTACAGAA 5079
Db 5176 AACGGGTCTGTGATCTGTGAGCCATGCCAGAGAAAGTGTGTCTCTCCGCTACAGAA 5235
QY 5080 AACCTCAGCAAAATCTGATCCGGAAGAGATAGAGACCTCAGAGCCCTGCGAGCTGTATC 5139
Db 5236 AACCTCAGCAAAATCTGATCCGGAAGAGATAGAGACCTCAGAGCCCTGCGAGCTGTATC 5295
QY 5140 CACTTCAACCAATTACAGTATCTCATTTGGAACCAATTAATTTCTACGAAATCGACATGAAG 5199
Db 5296 CACTTCAACCAATTACAGTATCTCATTTGGAACCAATTAATTTCTACGAAATCGACATGAAG 5355
QY 5200 CAGTACAGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGGCAGCTGTGTGTT 5259
Db 5356 CAGTACAGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGGCAGCTGTGTGTT 5415
QY 5260 GCCGCCTCTTCCAAACAGCTTCCCTGTCTCAATCTGCAAGTGAACAGCGAGGCGAGCGA 5319
Db 5416 GCCGCCTCTTCCAAACAGCTTCCCTGTCTCAATCTGCAAGTGAACAGCGAGGCGAGCGA 5475
QY 5320 GAGGAGTACTTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGGAATTTTACGGAAGAGCT 5379
Db 5476 GAGGAGTACTTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGGAATTTTACGGAAGAGCT 5535
QY 5380 AGCCGCAAGAGCTCTCAAGTGGAGTCCGTTACCTTTGGCCTTTGGCTACAGAGAACCC 5439
Db 5536 AGCCGCAAGAGCTCTCAAGTGGAGTCCGTTACCTTTGGCCTTTGGCTACAGAGAACCC 5595
QY 5440 TATCTGTTTGTGACCACTTCAACTCACTCAAGTAAATGAGATCCAGGACCGCTCTCA 5499
Db 5596 TATCTGTTTGTGACCACTTCAACTCACTCAAGTAAATGAGATCCAGGACCGCTCTCA 5655
QY 5500 GCAGGACCCCTGCGCAGAGGTACTCGACATCCGAAACCGCGCTACCTGGGCGCTGCC 5559
Db 5656 GCAGGACCCCTGCGCAGAGGTACTCGACATCCGAAACCGCGCTACCTGGGCGCTGCC 5715


```
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCAGGACAGCTCCTCAGCA 5502
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCAGGACAGCTCCTCAGCA 5652
QY 5503 GGGACCCCTGCGGAGCGTACCTGACATCCGAAACCGCGCTACTCGGGCCCTGCAATT 5562
Db 5653 GGGACCCCTGCGGAGCGTACCTGACATCCGAAACCGCGCTACTCGGGCCCTGCAATT 5712
QY 5563 TCCTCAGGAGCGATTTACTTTGGCGTCTCATACAGGATAAAATTAAAGGTCATTTGCTGC 5622
Db 5713 TCCTCAGGAGCGATTTACTTTGGCGTCTCATACAGGATAAAATTAAAGGTCATTTGCTGC 5772
QY 5623 AAGGGAACCTGTTGAAGAGTCCGCACTGAACACACCGGGCGCTGCACTTCCCGC 5682
Db 5773 AAGGGAACCTGTTGAAGAGTCCGCACTGAACACACCGGGCGCTGCACTTCCCGC 5832
QY 5683 AGCAGCCCAACAAAGCGAGCCACCCACGTCACAGGACATCACCAAGCGCTGACC 5742
Db 5833 AGCAGCCCAACAAAGCGAGCCACCCACGTCACAGGACATCACCAAGCGCTGACC 5892
QY 5743 TCAGCCCAAGCGCCGCGGAGCCGAGCCAGCCAGGAGAGCCAGCACCACCGC 5802
Db 5893 TCAGCCCAAGCGCCGCGGAGCCGAGCCAGGAGAGCCAGCACCACCGC 5952
QY 5803 TACCGGAGGCGGAGCCGAGCTGGCAGGACAGTCTCTGGCGGCGCCCTGAGCGA 5862
Db 5953 TACCGGAGGCGGAGCCGAGCTGGCAGGACAGTCTCTGGCGGCGCCCTGAGCGA 6012
QY 5863 GAGAGTCCCGCGCGGAGTGTCTAGCAGCGCGAGAGCGGTCCCGGGAGGCTGTTT 5922
Db 6013 GAGAGTCCCGCGCGGAGTGTCTAGCAGCGCGAGAGCGGTCCCGGGAGGCTGTTT 6072
QY 5923 GAAGCAGCAGAGGGGCGCGTGTCTGGGAGCGGTGAGGACCGCTGTCCAGGTG 5982
Db 6073 GAAGCAGCAGAGGGGCGCGTGTCTGGGAGCGGTGAGGACCGCTGTCCAGGTG 6132
QY 5983 AACAGGGAAGAGGCGAGAGTGCCTCTCAAGTTTTCAGGTTTACACTGTCACT 6037
Db 6133 AACAGGTTGGGACCAAGTCTTCAGTATTAATCTCAGCCAGGAAAAACCACTCCT 6187

RESULT 10
AX503778
LOCUS AX503778 6201 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 8 from Patent WO0226826.
ACCESSION AX503778
VERSION AX503778.1 GI:23385964
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gerlach,V.L., Macdougall,J.R., Smithson,G., Millet,I., Stone,D.,
Gunther,E., Ellerman,K., Grosse,W.M., Alsobrook,J.P., Lepley,D.M.,
Burgess,C.B., Padigar,M., Kekuda,R., Spytek,K.A., Leach,M.D. and
Shinkets,R.A.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0226826-A 8 04-APR-2002;
Curagen Corporation (US)
FEATURES
source
1. .6201
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.4%; Score 5629.8; DB 6; Length 6201;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;

QY 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATGCTGGTGTGTAACCCATT 60
```

```
Db 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATGCTGGTGTGTAACCCATT 60
QY 61 GCCAGCGGGGCTCCAGGCTGAATCTCTTCTCCAGGGGAAACACACCTTTATGACTCAA 120
Db 61 GCCAGCGGGGCTCCAGGCTGAATCTCTTCTCCAGGGGAAACACACCTTTATGACTCAA 120
QY 121 CAGCAGATGCTCTCTTTCCGAGAAAGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 121 CAGCAGATGCTCTCTTTCCGAGAAAGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCACTTTGTCGGAAGTAT 240
Db 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCACTTTGTCGGAAGTAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTTCGGCAAAGGACTTCGAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTTCGGCAAAGGACTTCGAAGTCAGA 300
QY 301 AGTCTTGTAGTGTGTGTCATTTGCTGAAGTGCAGTGTGTAAGAGAAGCAACCGGG 360
Db 301 AGTCTTGTAGTGTGTGTCATTTGCTGAAGTGCAGTGTGTAAGAGAAGCAACCGGG 360
QY 361 GACATCTATGCTATGAAGTGAAGAAAGAGGCTTTATTTGGCCCAAGGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGTGAAGAAAGAGGCTTTATTTGGCCCAAGGAGCAGGTTTCA 420
QY 421 TTTTGTGAGGAAGCGGAAACATATTATCTGAAGCAAAAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGGAAGCGGAAACATATTATCTGAAGCAAAAGCCCGTGGATCCCCCAATTA 480
QY 481 CAGTATGCCCTTTCAGGACAAATAATCCTTTATCTGATGGAGGAATATCAGCTTGGAGGG 540
Db 481 CAGTATGCCCTTTCAGGACAAATAATCCTTTATCTGATGGAGGAATATCAGCTTGGAGGG 540
QY 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACAGATTAGATGAAAACCTGATACAGTTT 600
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACAGATTAGATGAAAACCTGATACAGTTT 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGCTTCACTGATGGATACGTCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGCTTCACTGATGGATACGTCATCGA 660
QY 661 GACATCAAGCTTGAGAAACATTTCTGTTGACCGCACAGACACATCAAGCTTGGTGGATTTT 720
Db 661 GACATCAAGCTTGAGAAACATTTCTGTTGACCGCACAGACACATCAAGCTTGGTGGATTTT 720
QY 721 GGATCTCGCGGAAATGAATTCAAAACAGATGCTGATGCCAAACCTCCCGATTGGGACC 780
Db 721 GGATCTCGCGGAAATGAATTCAAAACAA---GGTGAATGCCAAACCTCCCGATTGGGACC 777
QY 781 CCAGATTACATGGCTCTGAAAGTGTGACCTGATGAACCGGGATGAAAGGACACCTAC 840
Db 778 CCAGATTACATGGCTCTGAAAGTGTGACCTGATGAACCGGGATGAAAGGACACCTAC 837
QY 841 GGCTGGAAGTGTGACTGCTGCTGCTGAGTGGGCGTGAATTCCTATGAGATGATTTATGGAGA 900
Db 838 GGCTGGAAGTGTGACTGCTGCTGCTGAGTGGGCGTGAATTCCTATGAGATGATTTATGGAGA 897
QY 901 TCCCGCTTCGACAGGGAACCTTGGCCAGAACTTCATTAACATTAATGATTTCCAGCGG 960
Db 898 TCCCGCTTCGACAGGGAACCTTGGCCAGAACTTCATTAACATTAATGATTTCCAGCGG 957
QY 961 TTTTGTGAAATTTTCAGATGACCCCAAGTGAGCAGTGAATTTCTTGTATCTGATTCAAAGC 1020
Db 958 TTTTGTGAAATTTTCAGATGACCCCAAGTGAGCAGTGAATTTCTTGTATCTGATTCAAAGC 1017
QY 1021 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCGCAATCCTTTCTTC 1080
Db 1018 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCGCAATCCTTTCTTC 1077
QY 1081 TCTAAATTTGACTGGAACAACTTCTGTAATCTCTCCCTCCCTTCTGTTCCCAACCTCAAG 1140
```

1078 TCTAAATTAAGTGAACAAACATTCGTAAAGCTCTCTCCCTCCCTTCTGTTCCCTCCCTCAAG 1137
1141 TCTGACGATGACACCTCCCAATTTTGTATGAACACAGAGAAATTCGTGGGTTTTCATCCTCT 1200
1138 TCTGACGATGACACCTCCCAATTTTGTATGAACACAGAGAAATTCGTGGGTTTTCATCCTCT 1197
1201 CCGTGCAGCTGAGCCCTCAGCTTCTCGGGTGAAGAACTCGCCGTTTGTGGGGTTTTCG 1260
1198 CCGTGCAGCTGAGCCCTCAGCTTCTCGGGTGAAGAACTCGCCGTTTGTGGGGTTTTCG 1257
1261 TACAGCAAGGCACTGGGGAATCTTGTGTAGATCTGAGTCTGTTGTGTCGGGTCTGAGCTCC 1320
1258 TACAGCAAGGCACTGGGGAATCTTGTGTAGATCTGAGTCTGTTGTGTCGGGTCTGAGCTCC 1317
1321 CCGTGCAGCTGAGCTTCTGAGTGAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
1318 CCGTGCAGCTGAGCTTCTGAGTGAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1377
1381 TCTCAGGCAAGGTGTCACAGATGGAGCGAAGTACCGGTTACATCGGAGGTGCA 1440
1378 TCTCAGGCAAGGTGTCACAGATGGAGCGAAGTACCGGTTACATCGGAGGTGCA 1437
1441 GAGGTGGAGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
1438 GAGGTGGAGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497
1501 TCCCTCTGGAGCAGGACTTGTACTATCATACAGATGAGTAGCTTAAAGCGAAGT 1560
1498 TCCCTCTGGAGCAGGACTTGTACTATCATACAGATGAGTAGCTTAAAGCGAAGT 1557
1561 TTGAGCAAGCAGGATGAGGTGTCAGAGGAGGTGAGCTGAAGGCTCTGAGACTTCCAT 1620
1558 TTGAGCAAGCAGGATGAGGTGTCAGAGGAGGTGAGCTGAAGGCTCTGAGACTTCCAT 1617
1621 GATATCAGAGCAGAGCGGAGCTCCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680
1618 GATATCAGAGCAGAGCGGAGCTCCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677
1681 GTGGAAGAAATGAGGTGATGATGAATCACTGGAAGGATCTGTCTCAGCAAGAGA 1740
1678 GTGGAAGAAATGAGGTGATGATGAATCACTGGAAGGATCTGTCTCAGCAAGAGA 1737
1741 CGGAGTATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGTCTGAGAAATTCAG 1800
1738 CGGAGTATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGTCTGAGAAATTCAG 1797
1801 CGGAAGCAGCAATGTCAGCAATTAATCTGTAAGGCTTAAGATCAAGGAGCTGAA 1860
1798 CGGAAGCAGCAATGTCAGCAATTAATCTGTAAGGCTTAAGATCAAGGAGCTGAA 1857
1861 GTGGGAGATATCGGAAATCTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
1858 GTGGGAGATATCGGAAATCTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1917
1921 CTCAGAGAGAACTGGAGAGGCT----- 1944
1918 CTCAGAGAGAACTGGAGAGGCTGTAAAGCCAGCAGCGAGGCCACCGAGCTGTGCGAG 1977
1945 -----GCAAGAGCGAGCCGAGAGGAGCTGAGAGGCTGAGAGCTGAGAGCCGAGAG 1992
1978 AATATCGGCAGCAAGAGCGAGCCGAGAGGAGCTGAGAGGCTGAGAGCCGAGAG 2037
1993 GATTTCTGAGGCAATCAAGAGAGCTGTTGGAAGCTGAGGAACGCCCAATTCCTG 2052
2038 GATTTCTTGAAGGCAATCAAGAGAGCTGTTGGAAGCTGAGGAACGCCCAATTCCTG 2097
2053 GAGAACAGGTGAAGAGCTTAGAGCAATGGAAGCTTAGAGAAACAGACTGAAGATGAC 2112
2098 GAGAACAGGTGAAGAGCTTAGAGCAATGGAAGCTTAGAGAAACAGACTGAAGATGAC 2157
2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCAAGAG 2172
2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCAAGAG 2217

2173 AAACATCGGAGGCCCAAGTCTCAGCCAGCACTAGAACTGCACTGAAACAGAAAGAG 2232
2218 AAACATCGGAGGCCCAAGTCTCAGCCAGCACTAGAACTGCACTGAAACAGAAAGAG 2277
2233 CAGCACTATGAGGAAGATTAAGTCTGGAACAATCAGATTAAGAAAGACCTGGCTGAC 2292
2278 CAGCACTATGAGGAAGATTAAGTCTGGAACAATCAGATTAAGAAAGACCTGGCTGAC 2337
2293 AAGGAGACATCTGGAGAACATGATGACAGAGACACAGGAGGAGGCCCATGAGAGGGCAAA 2352
2338 AAGGAGACATCTGGAGAACATGATGACAGAGACACAGGAGGAGGCCCATGAGAGGGCAAA 2397
2353 ATTCTCAGCGAACAGAAAGGCGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGGAA 2412
2398 ATTCTCAGCGAACAGAAAGGCGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGGAA 2457
2413 CAGAGGATTTGGAATCTGCTGAAGCCCAATAAACTTGCAGCAATAGCAGTCTTTTACC 2472
2458 CAGAGGATTTGGAATCTGCTGAAGCCCAATAAACTTGCAGCAATAGCAGTCTTTTACC 2517
2473 CAAAGGAACATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
2518 CAAAGGAACATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577
2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAG 2592
2578 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAG 2637
2593 AAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGCTGGAACCTGGAGACAAGATTG 2652
2638 AAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGCTGGAACCTGGAGACAAGATTG 2697
2653 CGGAGGCTCAGTCTAGAGCAGGAGGAGCAAACTGGAGCTCAAGCCAGCTCAGAGAG 2712
2698 CGGAGGCTCAGTCTAGAGCAGGAGGAGCAAACTGGAGCTCAAGCCAGCTCAGAGAG 2757
2713 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGCGAGGCTGACCGGCG 2772
2758 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGCGAGGCTGACCGGCG 2817
2773 GCCTTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGCAGAGCT 2832
2818 GCCTTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGCAGAGCT 2877
2833 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGATGCT 2892
2878 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGATGCT 2937
2893 CTTCTGAACAGCTGTACTGTAAATCAAGACTGGAAGAGCAGCTAAACAGAGCTGACGAG 2952
2938 CTTCTGAACAGCTGTACTGTAAATCAAGACTGGAAGAGCAGCTAAACAGAGCTGACGAG 2997
2953 GACAACTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3012
2998 GACAACTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3057
3013 GGCGCCCAACAGAGATTTGTAACCTGCGAAGTGAAGTGAGCAATCTCCCGGGAGATC 3072
3058 GGCGCCCAACAGAGATTTGTAACCTGCGAAGTGAAGTGAGCAATCTCCCGGGAGATC 3117
3073 ACGGAAACAGAGATGCGATTCAGCCAGAACAGATGAGGCTCTGAGAGACAG 3132
3118 ACGGAAACAGAGATGCGATTCAGCCAGAACAGATGAGGCTCTGAGAGACAG 3177
3133 TGCAACCATGCTGGAGGAACAGGTCATGGAATTTGGAGGCCCTAAACAGATGAGCTGAGAA 3192
3178 TGCAACCATGCTGGAGGAACAGGTCATGGAATTTGGAGGCCCTAAACAGATGAGCTGAGAA 3237
3193 AAAGAGCGGAGTGGGAGGCTGAGAGAGCTCTGGGTGATAGAAATCCGAGTTTGGAG 3252
3238 AAAGAGCGGAGTGGGAGGCTGAGAGAGCTCTGGGTGATAGAAATCCGAGTTTGGAG 3297

QY	3253	TGTCGGGTTGAGAGCTGCGAGAAATGCTGGACACCAGAAACAGAGCAGGCGGAGAGCC	3312	Db	4333	ACACACTTACCAGGCGCTTCTGCGGTGACAAATGAACCTCCCGAGTCTCCAGACCAAG	4392
Db	3298	TGTCGGGTTGAGAGCTGCGAGAGGATGCTGGACACCGAGAAACAGAGCAGGCGGAGAGCC	3357	QY	4393	GAGCCAGCAGCAGCTTGACACTTGGAAGGCTGATGAAGTGCCTCCAGGATTAACAACGA	4452
QY	3313	GATCAGCGGATCACCGAGTCTGCCAGGTGCTGGAGCTGGCAGTGAAGGACACAGGCT	3372	Db	4393	GAGCCAGCAGCAGCTTGACACTTGGAAGGCTGATGAAGTGCCTCCAGGATTAACAACGA	4452
Db	3358	GATCAGCGGATCACCGAGTCTGCCAGGTGCTGGAGCTGGCAGTGAAGGACACAGGCT	3417	QY	4453	GGACAGCAAGGCTGGGACAGAAAGTACATTGTCTCTGGAGGGATCAAAAGTCCTCATTTAT	4512
QY	3373	GAGATTCTCGCTCTGCGAGCAGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT	3432	Db	4453	GGACAGCAAGGCTGGGACAGAAAGTACATTGTCTCTGGAGGGATCAAAAGTCCTCATTTAT	4512
Db	3418	GAGATTCTCGCTCTGCGAGCAGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT	3477	QY	4513	GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCCTCCC	4572
QY	3433	GACAAAGCTCAATGACTGCGAGAGAGCAATGCTGCTTGAATGAATGCCCGAGCTTA	3492	Db	4513	GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCCTCCC	4572
Db	3478	GACAAAGCTCAATGACTGCGAGAGAGCAATGCTGCTTGAATGAATGCCCGAGCTTA	3537	QY	4573	GACGGGATGTATCTATTTCATGTGCGGTGGTCTCCGAACTCGCAATACAGCCAAA	4632
QY	3493	CAGCAGAGCTGGAGACTGAACAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA	3552	Db	4573	GACGGGATGTATCTATTTCATGTGCGGTGGTCTCCGAACTCGCAATACAGCCAAA	4632
Db	3538	CAGCAGAGCTGGAGACTGAACAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA	3597	QY	4633	GCA-----	4635
QY	3553	TTACAGCAGCAGATGGAAGCTGCGAGAAATACATTTTCCGTGAGCTCAAGGACTGCAA	3612	Db	4633	GCAGATGCCATACATACATGTAAGATGGAATCTCACCCGACACACCCTGCTGCCCGGG	4692
Db	3598	TTACAGCAGCAGATGGAAGCTGCGAGAAATACATTTTCCGTGAGCTCAAGGACTGCAA	3657	QY	4636	-----	4635
QY	3613	GAAGCTCTAGATCGGGCTGATCTACTGAAGCAGAAAGTGAAGTATCAGCTG	3672	Db	4693	AGAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGGCTTA	4752
Db	3658	GAAGCTCTAGATCGGGCTGATCTACTGAAGCAGAAAGTGAAGTATCAGCTG	3717	QY	4636	-----	4662
QY	3673	GAAAAATTCAGTCTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTCTCAA	3732	Db	4753	GAATCAGTTTCGAGGTGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGTAAACTG	4812
Db	3718	GAAPACATTCAGTCTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTCTCAA	3777	QY	4663	CTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTTAGACATGAACTGCACGCTG	4722
QY	3733	CAACCAAACTCATTTGATTTTCTCAAGCCAAATGGAACCACTGCTAAAAAGAAAG	3792	Db	4813	CTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTTAGACATGAACTGCACGCTG	4872
Db	3778	CAACCAAACTCATTTGATTTTCTCAAGCCAAATGGAACCACTGCTAAAAAGAAAG	3835	QY	4723	CCCTTCAGTGACCAAGTGTGTGTTGGGGCACCAGAGGAAGGCTCTACGCCCTGAATGC	4782
QY	3793	GGTTTATTAGTCGACGGAAGAGGACCTGTTTACCACACAGGTTTCTCTGAGTAC	3852	Db	4873	CCCTTCAGTGACCAAGTGTGTGTTGGGGCACCAGAGGAAGGCTCTACGCCCTGAATGC	4932
Db	3836	-----	3852	QY	4783	TTGAAAACTCCCTAACCCATGTCCAGGAAATTTGGAGCAGTCTTCCAAATTTATATATC	4842
QY	3853	AATGAGCTGAAGCTGGCCTGAGAGAGGAGAAAGTCTGTCGAGAGCTAGAGAAAGCC	3912	Db	4933	TTGAAAACTCCCTAACCCATGTCCAGGAAATTTGGAGCAGTCTTCCAAATTTATATATC	4992
Db	3853	AATGAGCTGAAGCTGGCCTGAGAGAGGAGAAAGTCTGTCGAGAGCTAGAGAAAGCC	3912	QY	4843	AAGGACCTGGAGAGCTTACTCATGATAGCAGGAGAGAGCGGGCACTGTGTCTTTGGAC	4902
QY	3913	CTTCAGAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAGCTGCCACCGCAAGCA	3972	Db	4993	AAGGACCTGGAGAGCTTACTCATGATAGCAGGAGAGAGCGGGCACTGTGTCTTTGGAC	5052
Db	3913	CTTCAGAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAGCTGCCACCGCAAGCA	3972	QY	4903	GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCGCGCATCTCA	4962
QY	3973	ACGGACCCACACACCCATCCAGCCAGCCACCGCGAGGAGCAGATGCCATGTCCGCC	4032	Db	5053	GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCGCGCATCTCA	5112
Db	3973	ACGGACCCACACACCCATCCAGCCAGCCACCGCGAGGAGCAGATGCCATGTCTGCC	4032	QY	4963	CCCAACATTTTGAAGCTCTCAAGGGCTGCCACTTGTGTTGGGCGAGGAGATTGAGAAC	5022
QY	4033	ATCGTCCGTCGCGAGAGCACAGCCAGTGCATGAGCTGTGCGCCCGCGCATCCAGC	4092	Db	5113	CCCAACATTTTGAAGCTCTCAAGGGCTGCCACTTGTGTTGGGCGAGGAGATTGAGAAC	5172
Db	4033	ATCGTCCGTCGCGAGAGCACAGCCAGTGCATGAGCTGTGCGCCCGCGCATCCAGC	4092	QY	5023	GGGCTCTGCAATCTGTGCAGCCTCCAGCAAGTCTCATTTCTCCGCTACCAACGAAAC	5082
QY	4093	CCGAGAAAGAGCTTCAACTCCAGAGAAATTTAGTCCGGCTTTAAGGAAACGATGCAC	4152	Db	5173	GGGCTCTGCAATCTGTGCAGCCTCCAGCAAGTCTCATTTCTCCGCTACCAACGAAAC	5232
Db	4093	CCGAGAAAGAGCTTCAACTCCAGAGAAATTTAGTCCGGCTTTAAGGAAACGATGCAC	4152	QY	5083	CTCAGCAATACTGCAATCCGGAAGAGATAGAGCCTCAGAGCCCTGAGCTGTATCCAC	5142
QY	4153	CACAATATTCCTCAACGATTAAGTGAATGGAATGGAATGGAATGGAATGGAATGGA	4212	Db	5233	CTCAGCAATACTGCAATCCGGAAGAGATAGAGCCTCAGAGCCCTGAGCTGTATCCAC	5292
Db	4153	CACAATATTCCTCAACGATTAAGTGAATGGAATGGAATGGAATGGAATGGAATGGA	4212	QY	5143	TTCAACCAATTTACAGTATCTCTCATTGGAAACCAATTAATTTCTACGAAATCGACATGAAGCAG	5202
QY	4213	TGCTGTGATACCGTCTGACCTTTGAGCGCCAGGATCCAAATGCTCAGATGTGAGTGATG	4272	Db	5293	TTCAACCAATTTACAGTATCTCTCATTGGAAACCAATTAATTTCTACGAAATCGACATGAAGCAG	5352
Db	4213	TGCTGTGATACCGTCTGACCTTTGAGCGCCAGGATCCAAATGCTCAGATGTGAGTGATG	4272	QY	5203	TACAGCTCGAGGAATTCCTTGATTAAGATGACCATTCCTTGGCAGCTGCTGTGTTGCC	5262
QY	4273	TGTACCCCAAGTGTCCAGCTGCTTGGCAGCCACCTGCGGTGCTGCTGCTGAATATGCC	4332	Db	5353	TACAGCTCGAGGAATTCCTTGATTAAGATGACCATTCCTTGGCAGCTGCTGTGTTGCC	5412
Db	4273	TGTACCCCAAGTGTCCAGCTGCTTGGCAGCCACCTGCGGTGCTGCTGCTGAATATGCC	4332	QY	5263	GCCTCTTCCACAGCTTCCCTGTCTCAATGCTGAGTGAAACGCGCAGGCGAGGAG	5322
QY	4333	ACACACTTACCAGGCGCTTCTGCGGTGACAAATGAATGCCCTCCAGGCTCTCCAGACCAAG	4392				

Db	5413	GCCTCTTCCACAGCTTCCTCTCTCAATCGTCAGGTGAACAGCGCAGCGCAGAG	5472
Qy	5323	GAGTACTTCTGCTGTTTCCACGAATTTGAGTGTCTGCTGATTTCTACGAGAGCTAGC	5382
Db	5473	GAGTACTTCTGCTGTTTCCACGAATTTGAGTGTCTGCTGATTTCTACGAGAGCTAGC	5532
Qy	5383	CGACAGAGCATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT	5442
Db	5533	CGACAGAGCATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT	5592
Qy	5443	CTGTTTGTGACCCACTCACTCACTGAGTAAATGAGATCCAGGACGCTCTCAGCA	5502
Db	5593	CTGTTTGTGACCCACTCACTCACTGAGTAAATGAGATCCAGGACGCTCTCAGCA	5652
Qy	5503	GGGACCCCTGCGCGAGCGGTACCTGGACATCCGGAACCCCGCGTACCTGGCGCCTGCCATT	5562
Db	5653	GGGACCCCTGCGCGAGCGGTACCTGGACATCCGGAACCCCGCGTACCTGGCGCCTGCCATT	5712
Qy	5563	TGCTCAGGAGCGATTTACTTGGGTCTCTATACGAGGATAAATTAAGGTCTATTGCTGC	5622
Db	5713	TGCTCAGGAGCGATTTACTTGGGTCTCTATACGAGGATAAATTAAGGTCTATTGCTGC	5772
Qy	5623	AAGGGAAACCTCTGTAAGGAGTCCGGCACTGAACACACCGCGGCGCGTCCACCTCCCGC	5682
Db	5773	AAGGGAAACCTCTGTAAGGAGTCCGGCACTGAACACACCGCGGCGCGTCCACCTCCCGC	5832
Qy	5683	AGCAGCCCCAACAGCGAGCCCAACCACTGTAACAGGACATCAACAGCGCTGCGCC	5742
Db	5833	AGCAGCCCCAACAGCGAGCCCAACCACTGTAACAGGACATCAACAGCGCTGCGCC	5892
Qy	5743	TCAGCGCCAGCGCGCGCCGAGGCGCCAGCCACCGCGGAGCCAGCACACCCACCGC	5802
Db	5893	TCAGCGCCAGCGCGCGCCGAGGCGCCAGCCACCGCGGAGCCAGCACACCCACCGC	5952
Qy	5803	TACCGCGAGCG	5862
Db	5953	TACCGCGAGCG	6012
Qy	5863	GAGAGTCCCG	5922
Db	6013	GAGAGTCCCG	6072
Qy	5923	GAGAGCAGCAGCG	5982
Db	6073	GAGAGCAGCAGCG	6132
Qy	5983	AACAGGGAAGAGGGA 5999	
Db	6133	AACAGGGAAGAGGGA 6149	

RESULT 11
AY257469
LOCUS
DEFINITION
AY257469 Homo sapiens rho/rac-interacting citron kinase (CIT) mRNA, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 8576)
Huang, C.Q., Wu, S.L., Shan, Y.X., Liu, S. and Xiao, P.J.
Direct Submission
Submitted (18-MAR-2003) Department of Biochemistry and Molecular
Biology, Basic Medicine, Suzhou University, Renming Road 49,
Suzhou, Jiangsu 215007, China
LOCATION/Qualifiers
1. .8576
/organism="Homo sapiens"
/mol_type="mRNA"

gene	1. .8576
CDS	54. .6137
db	taxon:9606
chromosome	"12"
map	"12q24.23"
gene	"CIT"
notes	"synonym: CRIK"
gene	"CIT"
notes	"serine/threonine protein kinases"
codon_start	1
product	"rho/rac-interacting citron kinase"
protein_id	"AAFI328.1"
db	taxon:9606
translation	"MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFTQQOMS PLREGILDALFVFECSQALMKHYNSFVKYSDTIAELQELQPSAKDFEVSRL VGGCHFAEVQVREKATGDIYAMKMKKALLAQEQVSFEERNILSRSTSWIPQL OYAFQKQHYLVMEYQPGGDLNRYEDOLDENLIQFYLAELILAVHSVLMGVY HRIKPNILVDTGHIKLVDGSAKMSNMVNAKLPIGTPDYNAPELVTVMGDG KTYGLDCWASVGVYAYEMIYGRSPFAGTSARTNNIMNFORFLKFPDDPKVSSDP LDLILQSLCCGKERLAFEGCLCHPFFSKIDWNINRSPFPFVILKSDDDTINFPDFE XNWSVSSPQSLSPSGFSGEELPFVGFYSKALGILGRSESVVSGJDSFAKTSMEKK LLIKSKELQDSQDKHMEQEMTRHRRYSVEVAVLSQKVELKASGTQRSLLEQLA TYITEGSLKRSLEQARMEVSQDDKALQLLHDIREQSRKLQIKIEQYQAOVEEVL MMNQLBEDLVARRSDLYSELRESRLAAEFKKAECQHLKAKQKQGEVEGY AKLEKINAGOLKIQLOKLEKAVKATEATELLONIQAKERAERELEKLNQNRDS SEGIRKLVEAELEKHEAQAQVSAHLEWHLKQKQHYEELIKVLDNQIKKLADKE TLENMORHEEAEHEKILISQKMINAMDSKIRLEQRIVELSEANKLANSSUFT QRNKAEEMISERQKQFYLEQAGLEQNRKLEQLEKISHQSDKNRLELEET RLREVEHEEQLEKRLQTELQLESQTLALQARALESQKRAKTELEET TAAEYIQAHTAHDEIQKFDALRNSCTVITDLEQNLQNTDNEALNNQVLYSK QLDEASGANDIQLRESEVDHLRREITEREMOLTSOKTMEALKTTCTMLEQVMDLE ALNDELLEKQWENRSLVDEKQFECVRELOHMLDTEKQSRADORLITEQV VELAHEKRAELIALQALKEQKKAESLSDKLNLEKHAMLENARSLQKLETER ELKQRLSEQAQKQMDQKXHIPLTQGLQALDRLKTERSDLSYQLENQVL YSEKVMGTTISQOTKLIDFQAKMDQPAKKKGLFSRRKEDPALPTVQPIQYNSLK LALKEKPAEALQKTRTELRSAREAAHRAKATPHSPSTATATQQAAMSAIK RSEPHQSAVSLILAPSSRRKSSSTPEFSRLKERMHNIPIHNFVNLNMRKACAV CLDTVHFGQAQSKLECCVMCHPKSTCLPATCGLPAAVATHTFAFCDKNSPGLQ TKZPSSSLHLEGMKVPNNKKGQCGQDRKIYVLEGSKULIYDNEAREAGQRPVEFE LCPLDGVSLHGVAGSELANTAKADVPYILKAMESPHPTCPGRTLYLLASFPDKQ RWYALDVSVAGVRSREKAEADKLLSLKLEGGDRDMDNCTLPFSDQVVLVGT EGYALNLSVHLVHVGAVFYIILKLEKLLMAGEBALCLVDVKKVQSLKIRK SHLPAPDTPNIFEAQVCHLFGAGKINGLICCAAMPKSVIILYNNELKCYIKR EITSPSCITHTNYTILIGTNKVEIDMKQVTEFLDKNDHSLAPFAVFAANSF PVSIVQNSAGOREEYLLCFHEFQVDSYGRSRSTDLLKWSRLPLAPAYRPLVPT HFNLSVIEIQASSAGTAPAYLDIPNRYLGPALSSGAIYLAASYODKLAVICCKG NLVKGSGTBHHRGPTSRSPNKRGPPTTNEHITKRVASSPAPPPGSPSPRSPHR YREGTELRDRKSPGRPLREKSPGRMLSTRERSPGRFLFSGRLPAGAVRTPPLS QVKNVWDQSSV"

ORIGIN

Query Match	89.9%	Score	5534.8	DB	9	Length	8576
Best Local Similarity	94.8%	Pred. No.	0				
Matches	5857	Conservative	0	Mismatches	107	Indels	228
Gaps	3						
Qy	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATTCCTTTGATGCTGGTCTGCTGCAACCCATT	60				
Db	54	ATGTTGAAGTTCAAAATATGAGCGCGGAATTCCTTTGATGCTGGTCTGCTGCAACCCATT	113				
Qy	61	GCCAGCGCGGCTCCAGGCTGATCTGTTCTCCAGGGGAACCCACCTTTTACTCTAA	120				
Db	114	GCCAGCGCGGCTCCAGGCTGAAATCTGTTCTCCAGGGGAACCCACCTTTTACTCTAA	173				
Qy	121	CACGAGATGTCCTCTCTTTCCCGAGAGGGATATTAGATGCCCTCTTTGCTTTTGA	180				
Db	174	CACGAGATGTCCTCTCTTTCCCGAGAGGGATATTAGATGCCCTCTTTGCTTTTGA	233				
Qy	181	GATGAGTCTGCTGCTGATGAGATTAAGCAGTGAACACTTTTCGCGGATAT	240				
Db	234	GAATGAGTCTGCTGCTGATGAGATTAAGCAGTGAACACTTTTCGCGGATAT	293				
Qy	241	TCCGACACCATAGTCTGAGTTACAGGAGCTCCAGCCCTTCGCGAAAGGACTTCGAAAGT	300				

Db	294	TCCGACACCATAGCTGAGTTTACAGSAGCTTCCAGCCTTCGGCAAAAGGACTTTCCGAAGTTCAGA	353
Qy	301	AGTCTTTGTAGGTTGTGGTTCACCTTTGTCTGAAGTGCAGGTGGTAAAGAGAGAACAACCGGG	360
Db	354	AGTCTTTGTAGGTTGTGGTTCACCTTTGTCTGAAGTGCAGGTGGTAAAGAGAGAACAACCGGG	413
Qy	361	GACATCTATGCTATGAAGAGTGATGAAGAGAGAGGCTTTATTTGGCCCCAGAGCAGGTTTCA	420
Db	414	GACATCTATGCTATGAAGAGTGATGAAGAGAGAGGCTTTATTTGGCCCCAGAGCAGGTTTCA	473
Qy	421	TTTTTTTGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATT	480
Db	474	TTTTTTTGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATT	533
Qy	481	CAGATGCGCTTTCAAGACAAAAATCACTTTATCTGATGGAGGAATATCAGCCTCGAGGG	540
Db	534	CAGTATGCTTTTCAGACAAAAATCACTTTATCTGCTCATGTGAAATATCAGCCTCGAGGG	593
Qy	541	GACTTGTGTTCATTTTGAATAGATAGAGACCAAGTTAGATGAAACCTGTATACAGTTT	600
Db	594	GACTTGTGTTCATTTTGAATAGATAGAGACCAAGTTAGATGAAACCTGTATACAGTTT	653
Qy	601	TACCTAGCTGAGCTGATTTTGGCTGTTTCAACAGCGTTTCATCTGATGGGATACGTGCATCGA	660
Db	654	TACCTAGCTGAGCTGATTTTGGCTGTTTCAACAGCGTTTCATCTGATGGGATACGTGCATCGA	713
Qy	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGTGGAATTT	720
Db	714	GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGTGGAATTT	773
Qy	721	GGATCTGCCGCAAAATGAAATTCAAACAAGATGGTGAATGCCAAATCCCGATTTGGGACC	780
Db	774	GGATCTGCCGCAAAATGAAATTCAAACAAGATGGTGAATGCCAAATCCCGATTTGGGACC	833
Qy	781	CCAGATTTACATGGCTCTGAAGTGTGACTGTGATGAAACGGGATGGAAAAGCACCTAC	840
Db	834	CCAGATTTACATGGCTCTGAAGTGTGACTGTGATGAAACGGGATGGAAAAGCACCTAC	893
Qy	841	GGCTGACTGTGACTGCTGGTTCAGTGGGGCGTGAATGCCCTATCAGATGATTTATGGGAGA	900
Db	894	GGCTGACTGTGACTGCTGGTTCAGTGGGGCGTGAATGCCCTATCAGATGATTTATGGGAGA	953
Qy	901	TCCCTCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATAACTATTGAATTTCCACGG	960
Db	954	TCCCTCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATAACTATTGAATTTCCACGG	1013
Qy	961	TTTTTTGAAATTTCCAGATGACCCCAAGTCAGAGTGCATTTCTTGATCTGATTCAAAGC	1020
Db	1014	TTTTTTGAAATTTCCAGATGACCCCAAGTCAGAGTGCATTTCTTGATCTGATTCAAAGC	1073
Qy	1021	TTGTTGTGCGGCCAGAAAGAGAGCTGAAGTTTGAAGGTTTGTGCTGCCATCTTTCTTC	1080
Db	1074	TTGTTGTGCGGCCAGAAAGAGAGCTGAAGTTTGAAGGTTTGTGCTGCCATCTTTCTTC	1133
Qy	1081	TCATAAATTTGACTGGAACAACATTCGTAACTCTCTCCCTCCCTCGTTCCACCCCTCAAG	1140
Db	1134	TCTAAAAATTTGACTGGAACAACATTCGTAACTCTCTCCCTCCCTCGTTCCACCCCTCAAG	1193
Qy	1141	TCTCAGCATGACACTCCAAATTTTGTGATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT	1200
Db	1194	TCTCAGCATGACACTCCAAATTTTGTGATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT	1253
Qy	1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTTCG	1260
Db	1254	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTTCG	1313
Qy	1261	TACAGCAAGGCACTGGGGATTTCTTGGTAGATCTCAGTCTGTGTGTGTCGGGTCTGGACTCC	1320
Db	1314	TACAGCAAGGCACTGGGGATTTCTTGGTAGATCTCAGTCTGTGTGTGTCGGGTCTGGACTCC	1373
Qy	1321	CCTGCCAAGACTAGCTCATCGGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1380
Db	1374	CCTGCCAAGACTAGCTCATCGGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1433

QY	1381	TCTCAGGACAAAGTGTCTCA	CAAGATGGAGCAGGAAATGA	CCCGGTTTACATCGGAGAGTGTCA	1444
DB	1434	TCTCAGGACAAAGTGTCTCA	CAAGATGGAGCAGGAAATGA	CCCGGTTTACATCGGAGAGTGTCA	1493
QY	1441	GAGGTGAGGCTGTGCTT	TAGTCAAGAGGAGGTGGAG	CTGAAGCCCTCTGAGACTCAGAGA	1500
DB	1494	GAGGTGAGGCTGTGCTT	TAGTCAAGAGGAGGTGGAG	CTGAAGCCCTCTGAGACTCAGAGA	1553
QY	1501	TCCCTCTCGGACGAGCA	CCCTTGTACTCATCATC	ACAGAATGCAGTAGCTTAAAGCGGAAGT	1560
DB	1554	TCCCTCTCGGACGAGCA	CCCTTGTACTCATCATC	ACAGAATGCAGTAGCTTAAAGCGGAAGT	1613
QY	1561	TTGGAGCAAGCAGGATGG	AGGTGTCCACGAGGATGCA	CAAAACACTGCAGCTTCTCCAT	1620
DB	1614	TTGGAGCAAGCAGGATGG	AGGTGTCCACGAGGATGCA	CAAAACACTGCAGCTTCTCCAT	1673
QY	1621	GATATCAGAGCAGACG	CCGGAAGCTCCAAAGAA	TCAAAGACGAGGAGTACCAGGCTCAA	1680
DB	1674	GATATCAGAGCAGACG	CCGGAAGCTCCAAAGAA	TCAAAGACGAGGAGTACCAGGCTCAA	1733
QY	1681	GTGGAAGAAATGAGTT	GTATGATCAATCAGTT	GGAGAGGATCTTGCTCTCAGCAAGAGA	1740
DB	1734	GTGGAAGAAATGAGTT	GTATGATCAATCAGTT	GGAGAGGATCTTGCTCTCAGCAAGAGA	1793
QY	1741	CGGAGTCAATCTCTAC	GAAATCTGAGCTCAGAG	AGTCTCGGCTTGCTCTGAAGAAATTC	1800
DB	1794	CGGAGTCAATCTCTAC	GAAATCTGAGCTCAGAG	AGTCTCGGCTTGCTCTGAAGAAATTC	1853
QY	1801	CGGAAAGCGACAGAA	TGTACGATAAACTGT	TGAAGCTAAGNATCAAGGGAGCCGTGA	1860
DB	1854	CGGAAAGCGACAGAA	TGTACGATAAACTGT	TGAAGCTAAGNATCAAGGGAGCCGTGA	1913
QY	1861	GTGGGAAATATCGAA	AACTGGAGAAGATCA	TGCTGAGCAGCAGCTCAAAATTCAGGAG	1920
DB	1914	GTGGGAAATATCGAA	AACTGGAGAAGATCA	TGCTGAGCAGCAGCTCAAAATTCAGGAG	1973
QY	1921	CTCCAAGAGAACTGG	AGAAAGCTGCAAAAG	GAGCGAGCGGAGAGCTGGAAGAGCTG	1980
DB	1974	CTCCAAGAGAACTGG	AGAAAGCTGCAAAAG	GAGCGAGCGGAGAGCTGGAAGAGCTG	2030
QY	1981	CAGAACCGAGAGGAT	CTTCTGAGGATCAG	AAAGAGCTGGTGGAGCTGAGCAAGCTG	2040
DB	2031	CAGAAATATCCGCC	AGGCAAGGACGAG	CCGAGAGGAGCTGGAAGAGCTGCAAA	2090
QY	2041	CGCATCTCTCGGAGA	CAAGGTAAAGAGA	CTAGAGACCATTGGAGCGGTAGAGAAAC	2100
DB	2091	GAGGATCTTCTGA	AGGCATCAGAAAG	AGCTGGTG-----	2126
QY	2101	CTGAAGGATGATCT	CCAGACAAAATCC	ACAGATCCAGCAGATGGCTGATAAAATCTG	2160
DB	2127	-----	-----	-----GAGGCTGAG	2135
QY	2161	GAGCTCGAAGAGAA	CAACTCGGAGGCC	CAAGTCTCAGCCACGACCTAGAAAGTGCA	2220
DB	2136	GAGCTCGAAGAGAA	CAACTCGGAGGCC	CAAGTCTCAGCCACGACCTAGAAAGTGCA	2195
QY	2221	AAACAGAAAGACG	ACGACTATGAGGAAA	AGATTAAAGTTTGGACAATCAGATAAAGAAA	2280
DB	2196	AAACAGAAAGACG	ACGACTATGAGGAAA	AGATTAAAGTTTGGACAATCAGATAAAGAAA	2255
QY	2281	GACCTGGCTGACAG	AGGAGACACTCGAG	ACATGATGACAGACACGAGGAGAGGCCCAT	2340
DB	2256	GACCTGGCTGACAG	AGGAGACACTCGAG	ACATGATGACAGACACGAGGAGAGGCCCAT	2315
QY	2341	GAGAAAGGCAAAAT	TTCTCAGCGAA	CAGAAAGCGATGATCAATGCTATGGATTCCA	2400
DB	2316	GAGAAAGGCAAAAT	TTCTCAGCGAA	CAGAAAGCGATGATCAATGCTATGGATTCCA	2375
QY	2401	AGATCCCTGGAA	CAGAGATTGTGG	AACTGTCTGAGCCCAATTAACCTTGACAGCAAA	2460
DB	2376	AGATCCCTGGAA	CAGAGATTGTGG	AACTGTCTGAGCCCAATTAACCTTGACAGCAAA	2435

QY	2461	AGTCTTTTATCCCAAAGAAACATGAAGCCCAAGAAAGAGATGATTTCTGAACCTCAGGCAA	2522
DB	2436	AGTCTTTTATCCCAAAGAAACATGAAGCCCAAGAAAGAGATGATTTCTGAACCTCAGGCAA	2495
QY	2521	CAGAAATTTTACCTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGAACCCGAAACTGGAG	2580
DB	2496	CAGAAATTTTACCTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGAACCCGAAACTGGAG	2555
QY	2581	GAGCAGCTTGGAGAAAGATCAGCCACCAAGACCAACAGTGAACAAGATTCGGCTGCTCGAACTG	2640
DB	2556	GAGCAGCTTGGAGAAAGATCAGCCACCAAGACCAACAGTGAACAAGATTCGGCTGCTCGAACTG	2615
QY	2641	GAGACAAAGATTCGGGAGGTCAGTCTTAGAGCACAGGAGCAGAACTTGGAGCTCAAGCGC	2700
DB	2616	GAGACAAAGATTCGGGAGGTCAGTCTTAGAGCACAGGAGCAGAACTTGGAGCTCAAGCGC	2675
QY	2701	CAGCTTCACAGAGCTACAGCTCTCCCTGTCAGGAGCGCAGTCAACAGTTGACAGCCCTGCAG	2760
DB	2676	CAGCTTCACAGAGCTACAGCTCTCCCTGTCAGGAGCGCAGTCAACAGTTGACAGCCCTGCAG	2735
QY	2761	GCTGCACGGGCGGCTTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTTGGAAAGAGACC	2820
DB	2736	GCTGCACGGGCGGCTTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTTGGAAAGAGACC	2795
QY	2821	ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCAGGGCACATAGAGATGAAATCCAGCGC	2880
DB	2796	ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCAGGGCACATAGAGATGAAATCCAGCGC	2855
QY	2881	AAATTTGATGCTCTTCGTAAACAGCTGACTGTATATACACAGACCTGGAGGAGCAGCTAAAC	2940
DB	2856	AAATTTGATGCTCTTCGTAAACAGCTGACTGTATATACACAGACCTGGAGGAGCAGCTAAAC	2915
QY	2941	CAGCTGACCGAGGACAAAGCTGAACTCAACAAACCAAACTTCTACTGTCTCCAAACAATC	3000
DB	2916	CAGCTGACCGAGGACAAAGCTGAACTCAACAAACCAAACTTCTACTGTCTCCAAACAATC	2975
QY	3001	GATGAGGCTTCTGGCGCCAAACGACGAGATTTGACACTGCGAAAGTGAAGTGGACCACTCTC	3060
DB	2976	GATGAGGCTTCTGGCGCCAAACGACGAGATTTGACACTGCGAAAGTGAAGTGGACCACTCTC	3035
QY	3061	CGCGCGAGATCACGGAACAGAGATGACCTTACAGCCAGAAAGCAACATGATGGAGGT	3120
DB	3036	CGCGCGAGATCACGGAACAGAGATGACCTTACAGCCAGAAAGCAACATGATGGAGGT	3095
QY	3121	CTGAAGACCACTGCAACATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAAACGAT	3180
DB	3096	CTGAAGACCACTGCAACATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAAACGAT	3155
QY	3181	GAGCTGCTAGAAAAAGACGCGCATGGAGGCTTGGAGAGCGCTCTCTGGGTGATGAGAAA	3240
DB	3156	GAGCTGCTAGAAAAAGACGCGCATGGAGGCTTGGAGAGCGCTCTCTGGGTGATGAGAAA	3215
QY	3241	TCCCAAGTTTCAGTGTCCGGTTCAGAGAGTCCAGAGAAATGCTGACACCCAGAAACAGAGC	3300
DB	3216	TCCCAAGTTTCAGTGTCCGGTTCAGAGAGTCCAGAGAAATGCTGACACCCAGAAACAGAGC	3275
QY	3301	AGGCGAGAGCCGATCAGCGGATTCACCGAGTCTCGCCAGGTGGTGGAGCTGCGAGTGAAG	3360
DB	3276	AGGCGAGAGCCGATCAGCGGATTCACCGAGTCTCGCCAGGTGGTGGAGCTGCGAGTGAAG	3335
QY	3361	GAGCACAAGGCTCAGATTTCTCGCTTCGACGAGGCTCTCAAGAGCAGAGCTGAGAGCC	3420
DB	3336	GAGCACAAGGCTCAGATTTCTCGCTTCGACGAGGCTCTCAAGAGCAGAGCTGAGAGCC	3395
QY	3421	GAGAGCCTCTTCACAAGCTCAATGACCTTGGAGAAAGCATGCTATGCTTGAATGAAT	3480
DB	3396	GAGAGCCTCTTCACAAGCTCAATGACCTTGGAGAAAGCATGCTATGCTTGAATGAAT	3455
QY	3481	GCCCGAAGCTTACAGCAGAGCTGGAGACTGACAGAGAGCTCAACAGAGGCTTCTCGAA	3540
DB	3456	GCCCGAAGCTTACAGCAGAGCTGGAGACTGACAGAGAGCTCAACAGAGGCTTCTCGAA	3515
QY	3541	GAGCAAGCCAAATTACAGCAGCAGATGGACCTGCGAGAAAAATCACATTTTTCGCTGACT	3600

3516	DB	GAGCAAGCCAAATTACAGCAGCAGATGGACCTGTCAGAAAAATCACATTTTCCGCTGTGACT	3575
3601	QY	CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGCAGAAAAAGTGAAGCTTG	3660
3576	DB	CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGCAGAAAAAGTGAAGCTTG	3635
3661	QY	GAGTATCAGCTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGGC	3720
3636	DB	GAGTATCAGCTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGGC	3695
3721	QY	ACTATTTCTCAACAAACGAACTCATTTGATTTTCTGCAAGCCAAAAATGHAACCAACTGCT	3780
3696	DB	ACTATTTCTCAACAAACGAACTCATTTGATTTTCTGCAAGCCAAAAATGHAACCAACTGCT	3755
3781	QY	AAAAAGAAAAAGGGTTTATTTAGTCGACGAAAGAGGACCTCTCTTTTACCACACACAGGTT	3840
3756	DB	AAAAAGAAAAAGGGTTTATTTAGTCGACGAAAGAGGACCTCTCTTTTACCACACACAGGTT	3815
3841	QY	CCTCTGCAGTACAATGAGCTGAAAGCTGGCCCTGGAGAAAGGAGAAAGCTCGCTGTGCAGAG	3900
3816	DB	CCTCTGCAGTACAATGAGCTGAAAGCTGGCCCTGGAGAAAGGAGAAAGCTCGCTGTGCAGAG	3875
3901	QY	CTAGAGAAAGCCCTTCAGAAAGACCGGCATCGAGCTCCGGTCCGCCGGAGGAGAGCTGCC	3960
3876	DB	CTAGAGAAAGCCCTTCAGAAAGACCGGCATCGAGCTCCGGTCCGCCGGAGGAGAGCTGCC	3935
3961	QY	CACGCAAAAGCAACGGACCAACCACACCCATCCACGCCAGCCAACCGAGGACGACAGATC	4020
3936	DB	CACGCAAAAGCAACGGACCAACCACACCCATCCACGCCAGCCAACCGAGGACGACAGATC	3995
4021	QY	GCCATGTCGCCATCGTGGCGGTGGCAGAGCACCAGCCAGTCCCATGAGCCCTGCTGGCC	4080
3996	DB	GCCATGTCGCCATCGTGGCGGTGGCAGAGCACCAGCCAGTCCCATGAGCCCTGCTGGCC	4055
4081	QY	CCGCCATCCAGCCGCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCTTAAG	4140
4056	DB	CCGCCATCCAGCCGCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCTTAAG	4115
4141	QY	GAAGCATGACACACAATATTCTCAACGATTCAACGTAGGACTGAAATGCGAGCCACA	4200
4116	DB	GAAGCATGACACACAATATTCTCAACGATTCAACGTAGGACTGAAATGCGAGCCACA	4175
4201	QY	AAGTGTCTGTGTCTGGATACCGTGCACTTTGGAGCGCCAGGCATCCAAATGTCTCGAA	4260
4176	DB	AAGTGTCTGTGTCTGGATACCGTGCACTTTGGAGCGCCAGGCATCCAAATGTCTCGAA	4235
4261	QY	TGTCAGGTGATGTGTACCCCAAGTGTCTCAGTGTTGCCAGCCACCTGCGGCTTGCT	4320
4236	DB	TGTCAGGTGATGTGTACCCCAAGTGTCTCAGTGTTGCCAGCCACCTGCGGCTTGCT	4295
4321	QY	GCTGAATATGCCACACATTCACCGAGGCTTCTGCCGTGACAAATGAACCTCCCCAGT	4380
4296	DB	GCTGAATATGCCACACATTCACCGAGGCTTCTGCCGTGACAAATGAACCTCCCCAGT	4355
4381	QY	CTCCAGACCAAGAGCCCGCAGCAGCTTGCACCTTGAAAGGGTGGATGAAGTGGCCAGG	4440
4356	DB	CTCCAGACCAAGAGCCCGCAGCAGCTTGCACCTTGAAAGGGTGGATGAAGTGGCCAGG	4415
4441	QY	AATAACAAACGAGCAGCAAGCTGGGACAGGAGTCAATTTCTCTCGAGGGATCAAAA	4500
4416	DB	AATAACAAACGAGCAGCAAGCTGGGACAGGAGTCAATTTCTCTCGAGGGATCAAAA	4475
4501	QY	GTCTCATTTATGCAATGAAGCCAGAGAGCTGGACAGAGCCCGGTGGAAAGATTTGAG	4560
4476	DB	GTCTCATTTATGCAATGAAGCCAGAGAGCTGGACAGAGCCCGGTGGAAAGATTTGAG	4535
4561	QY	CTGTGCTTTCCGACGGGATGATCTATTCTATGTTGCCGTGGTCTTCGAACTCGCA	4620
4536	DB	CTGTGCTTTCCGACGGGATGATCTATTCTATGTTGCCGTGGTCTTCGAACTCGCA	4595
4621	QY	AATACAGCCAAAAGCA-----	4635

QY 181 GAATGTCAGTCTGCTGATGAAGATTAAAGCACTGAGCAACTTTGTCCGAAAGTAT 240
Db 181 GAATGTCAGTCTGCTGATGAAGATTAAAGCACTGAGCAACTTTGTCCGAAAGTAT 240
QY 241 TCCGACACCATAGTGTAGTTACAGGAGCTCAGGCTTCGGCAAGAGACTTCGAAGTCAGA 300
Db 241 TCCGACACCATAGTGTAGTTACAGGAGCTCAGGCTTCGGCAAGAGACTTCGAAGTCAGA 300
QY 301 AGTCTTGTAGTGTGTGCTCACTTTGTGAAAGTGCAGGTGCTAAGAGAAAGCAACCGGG 360
Db 301 AGTCTTGTAGTGTGTGCTCACTTTGTGAAAGTGCAGGTGCTAAGAGAAAGCAACCGGG 360
QY 361 GACATCTATCTATGAAGTGTGAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 420
Db 361 GACATCTATCTATGAAGTGTGAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 420
QY 421 TTTTGTGAGGAAGCGGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGGAAGCGGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480
QY 481 CAGTATGCCCTTTTTCAGGACAAAATCACCTTTTATCTGATGAGGAAATATCAGCCTCGAGGG 540
Db 481 CAGTATGCCCTTTTTCAGGACAAAATCACCTTTTATCTGATGAGGAAATATCAGCCTCGAGGG 540
QY 541 GACTGCTGTCACTTTTGTATGATATGAGGACCACTTATGATGAGGAAATATCAGCCTCGAGGG 600
Db 541 GACTGCTGTCACTTTTGTATGATATGAGGACCACTTATGATGAGGAAATATCAGCCTCGAGGG 600
QY 601 TACTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTTCACTGATGGGATACCTGATCGA 660
Db 601 TACTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTTCACTGATGGGATACCTGATCGA 660
QY 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACATCAAGCTGGTGGATTTT 720
Db 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACATCAAGCTGGTGGATTTT 720
QY 721 GGATCTGCCGCGAAAATGAATTCAAACAGATGGTGAATGCCAAACTCCCGATTCGGACC 780
Db 721 GGATCTGCCGCGAAAATGAATTCAAACAGATGGTGAATGCCAAACTCCCGATTCGGACC 780
QY 781 CCAGATTACATGCTCCTGAAAGTGTGACTGTGATGAACCGGGATGGAAAGGCACTTAC 840
Db 781 CCAGATTACATGCTCCTGAAAGTGTGACTGTGATGAACCGGGATGGAAAGGCACTTAC 840
QY 841 GGGCTGGAATGTGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
Db 841 GGGCTGGAATGTGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
QY 901 TCCCTTCCGAGAGGAAACCTCTGCGAGAACCTTCAATTAACATTTATGAAATTCAGCGG 960
Db 901 TCCCTTCCGAGAGGAAACCTCTGCGAGAACCTTCAATTAACATTTATGAAATTCAGCGG 960
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGTGATGATGATGATGATGATGATGATG 1020
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGTGATGATGATGATGATGATGATGATG 1020
QY 1021 TTTTGTGCGGCGAGAAAGAGAGACTGAAGTTTGAAGTTCCTTGTGCTGCTTCTTCTTC 1080
Db 1021 TTTTGTGCGGCGAGAAAGAGAGACTGAAGTTTGAAGTTCCTTGTGCTGCTTCTTCTTC 1080
QY 1081 TCTAAATTTGATGTAACCAATTTGATGTAACCAATTTGATGTAACCAATTTGATGTAAC 1140
Db 1081 TCTAAATTTGATGTAACCAATTTGATGTAACCAATTTGATGTAACCAATTTGATGTAAC 1140
QY 1141 TCTGAGATGACACCTTCAATTTTGTGTAACCAATTTGATGTAACCAATTTGATGTAAC 1200
Db 1141 TCTGAGATGACACCTTCAATTTTGTGTAACCAATTTGATGTAACCAATTTGATGTAAC 1200
QY 1201 CCGTGCAGCTGAGCCCTTCCAGCTTCTCGGTGAAAGCACTGCCGTTTGTGGGTTTTCG 1260
Db 1201 CCGTGCAGCTGAGCCCTTCCAGCTTCTCGGTGAAAGCACTGCCGTTTGTGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGATTTCTTGTGATGATGATGATGATGATGATGATGATGATG 1320

Db 1261 TACAGCAAGGCACTGGGATTTCTTGTGATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 CTTGCAAGACTAGCTCCATGCAAAAGAAATCTTCTCATCAAAAGCAAAAGAGACTCAAGAC 1380
Db 1321 CTTGCAAGACTAGCTCCATGCAAAAGAAATCTTCTCATCAAAAGCAAAAGAGACTCAAGAC 1380
QY 1381 TCTCAGGCAAGTGTCAAGATGAGCAGAAATGACCGGTTTACATCGAGAGTGTCA 1440
Db 1381 TCTCAGGCAAGTGTCAAGATGAGCAGAAATGACCGGTTTACATCGAGAGTGTCA 1440
QY 1441 GAGTGGAGGCTGTCTTGTAGTCAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1441 GAGTGGAGGCTGTCTTGTAGTCAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500
QY 1501 TCCCTTCTGAGGAGGACCTTGTCTACCTACATCAGAAATGCAAGCACTGCGAGTTCCTCAT 1560
Db 1501 TCCCTTCTGAGGAGGACCTTGTCTACCTACATCAGAAATGCAAGCACTGCGAGTTCCTCAT 1560
QY 1561 TTGAGCAAGCACTGAGTGGAGTGTCTCCAGGAGATGCAAAAGCACTGCGAGTTCCTCAT 1620
Db 1561 TTGAGCAAGCACTGAGTGGAGTGTCTCCAGGAGATGCAAAAGCACTGCGAGTTCCTCAT 1620
QY 1621 GATATCAGAGAGCAGCGGAGCTCCAGAGCTCCAGAGATCAAAAGCACTGCGAGTTCCTCAT 1680
Db 1621 GATATCAGAGAGCAGCGGAGCTCCAGAGCTCCAGAGATCAAAAGCACTGCGAGTTCCTCAT 1680
QY 1681 GTGGAAGAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Db 1681 GTGGAAGAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGTGTCTGCGCTTGTCTGTAAGAAATCAAG 1800
Db 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGTGTCTGCGCTTGTCTGTAAGAAATCAAG 1800
QY 1801 CGGAAACGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1860
Db 1801 CGGAAACGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1860
QY 1861 GTGGAGAAATATCGAAATCTGAGCTGAGAGAGTCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
Db 1861 GTGGAGAAATATCGAAATCTGAGCTGAGAGAGTCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
QY 1921 CTCCAGAGAACTGGAGAGGCT----- 1944
Db 1921 CTCCAGAGAACTGGAGAGGCTTAAAGCCAGCAAGGAGGCTCCAGAGCTGTGAG 1980
QY 1945 -----GCAAGGAGGAGCGAGAGGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1992
Db 1945 -----GCAAGGAGGAGCGAGAGGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1992
QY 1993 GATTCCTTGAAGGCTCAGAAAGAGTGTGTAAGAGTGTGTAAGAGCTGAGAGAGCTGAGAG 2040
Db 1993 GATTCCTTGAAGGCTCAGAAAGAGTGTGTAAGAGTGTGTAAGAGCTGAGAGAGCTGAGAG 2040
QY 2041 GATTCCTTGAAGGCTCAGAAAGAGTGTGTAAGAGTGTGTAAGAGCTGAGAGAGCTGAGAG 2100
Db 2041 GATTCCTTGAAGGCTCAGAAAGAGTGTGTAAGAGTGTGTAAGAGCTGAGAGAGCTGAGAG 2100
QY 2053 GAGAAACAGAGTAAAGAGCTAGAGAGCCATGAGCGTGAAGAGAAACAGACTGAAGAGTAC 2112
Db 2053 GAGAAACAGAGTAAAGAGCTAGAGAGCCATGAGCGTGAAGAGAAACAGACTGAAGAGTAC 2112
QY 2101 GAGAAACAGAGTAAAGAGCTAGAGAGCCATGAGCGTGAAGAGAAACAGACTGAAGAGTAC 2160
Db 2101 GAGAAACAGAGTAAAGAGCTAGAGAGCCATGAGCGTGAAGAGAAACAGACTGAAGAGTAC 2160
QY 2113 ATCCAGCAAAATCCCAACAGATCCAGAGCTGAGTGTGATGATGATGATGATGATGATG 2172
Db 2113 ATCCAGCAAAATCCCAACAGATCCAGAGCTGAGTGTGATGATGATGATGATGATGATG 2172
QY 2161 ATCCAGCAAAATCCCAACAGATCCAGAGCTGAGTGTGATGATGATGATGATGATGATG 2220
Db 2161 ATCCAGCAAAATCCCAACAGATCCAGAGCTGAGTGTGATGATGATGATGATGATGATG 2220
QY 2173 AAACATCGGAGGAGCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAGAAACAGAGAG 2232
Db 2173 AAACATCGGAGGAGCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAGAAACAGAGAG 2232
QY 2221 AAACATCGGAGGAGCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAGAAACAGAGAG 2280
Db 2221 AAACATCGGAGGAGCCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAGAAACAGAGAG 2280
QY 2233 CAGCACTATGAGAAAGATTTAAAGTGTGGAATTCAGATGATGATGATGATGATGATGATG 2292
Db 2233 CAGCACTATGAGAAAGATTTAAAGTGTGGAATTCAGATGATGATGATGATGATGATGATG 2292
QY 2281 CAGCACTATGAGAAAGATTTAAAGTGTGGAATTCAGATGATGATGATGATGATGATGATG 2340
Db 2281 CAGCACTATGAGAAAGATTTAAAGTGTGGAATTCAGATGATGATGATGATGATGATGATG 2340
QY 2293 AAGGAGCACTGGAGAGATGATGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2352
Db 2293 AAGGAGCACTGGAGAGATGATGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2352

2341 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGAGGCCCATGAGAAGGSCAAA 2400 Db
2353 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGGATTCMAAGATCAGATCCCTGGAA 2412 QY
2401 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGGATTCMAAGATCAGATCCCTGGAA 2460 Db
2413 CAGAGATTGTGGAACTGTCTGAAGCCAATAAATTCAGCAAAATAGCAGTCTTTTACC 2472 QY
2461 CAGAGATTGTGGAACTGTCTGAAGCCAATAAATTCAGCAAAATAGCAGTCTTTTACC 2520 Db
2473 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCCAAAGAAAATTTTAC 2532 QY
2521 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCCAAAGAAAATTTTAC 2580 Db
2533 CTGGAGACACAGGCTCGGAAGTTGGAGGCCAGAACCCGAAAATCTGGAGGAGCAGCTGGAG 2592 QY
2581 CTGGAGACACAGGCTCGGAAGTTGGAGGCCAGAACCCGAAAATCTGGAGGAGCAGCTGGAG 2640 Db
2593 AAGATCAGCCACCAAGACCAACAGTGAAGAATTCGGCTCTGCGAACTGGAGACAAAGATTG 2652 QY
2641 AAGATCAGCCACCAAGACCAACAGTGAAGAATTCGGCTCTGCGAACTGGAGACAAAGATTG 2700 Db
2653 CGGGAGTCTAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGGCCAGCTCACAGAG 2712 QY
2701 CGGGAGTCTAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGGCCAGCTCACAGAG 2760 Db
2713 CTACAGTCTCTCCCTGCAGGAGCGCAGTCAACAGTTGACAGCCCTGCAGGCTGCACGGGCG 2772 QY
2761 CTACAGTCTCTCCCTGCAGGAGCGCAGTCAACAGTTGACAGCCCTGCAGGCTGCACGGGCG 2820 Db
2773 GCCCTGGAGACCCAGCTTCGCCAGGCGGAAGACAGAGCTGGAGAGACACACAGAGAGCT 2832 QY
2821 GCCCTGGAGACCCAGCTTCGCCAGGCGGAAGACAGAGCTGGAGAGACACACAGAGAGCT 2880 Db
2833 GAAGAGGAGATCCAGGACCTCAGCGGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2892 QY
2881 GAAGAGGAGATCCAGGACCTCAGCGGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940 Db
2893 CTTCTGTAACAGTGTACTGTATCAAGACCTGGAGGAGCAGCTAAACAGAGTGCACCGAG 2952 QY
2941 CTTCTGTAACAGTGTACTGTATCAAGACCTGGAGGAGCAGCTAAACAGAGTGCACCGAG 3000 Db
2953 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3012 QY
3001 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060 Db
3013 GCGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3072 QY
3061 GCGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3120 Db
3073 ACGGAACGAGAGATGCAGCTTACCAGCCAGAGACAAACGATGGAGGCTCTGAAGACCCAG 3132 QY
3121 ACGGAACGAGAGATGCAGCTTACCAGCCAGAGACAAACGATGGAGGCTCTGAAGACCCAG 3180 Db
3133 TGCACCATCTCGAGGAAACAGTGCATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3192 QY
3181 TGCACCATCTCGAGGAAACAGTGCATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240 Db
3193 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGATTTGAG 3252 QY
3241 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGATTTGAG 3300 Db
3253 TGTCCGGTTTCGAGAGCTGCAGAGATGCTGGACACCGGAAACAGACAGAGGCGGAGAGCC 3312 QY
3301 TGTCCGGTTTCGAGAGCTGCAGAGATGCTGGACACCGGAAACAGACAGAGGCGGAGAGCC 3360 Db
3313 GATCAGCGGATCACCGAGTCTCCAGGTGGTGGAGCTGGAGTGGAGTGAAGAGAGCACAAGGCT 3372 QY
3361 GATCAGCGGATCACCGAGTCTCCAGGTGGTGGAGCTGGAGTGGAGTGAAGAGAGCACAAGGCT 3420 Db
3373 GAGATTCTCGCTCTCAGCAGGCTCTCAAAGACGAGAGAGCTGAAGCCGAGAGCTCTCTCT 3432 QY
3421 GAGATTCTCGCTCTCAGCAGGCTCTCAAAGACGAGAGAGCTGAAGCCGAGAGCTCTCTCT 3480 Db

3433 GACAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCGAGCTTA 3492 QY
3481 GACAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCGAGCTTA 3540 Db
3493 CAGCAGAAGCTGGAGAGCTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAA 3552 QY
3541 CAGCAGAAGCTGGAGAGCTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAA 3600 Db
3553 TTACAGCAGAGATGGAGCCTGCAGAAAAATCACAATTTTCGCTCTGACTCAAGGACTGCA 3612 QY
3601 TTACAGCAGAGATGGAGCCTGCAGAAAAATCACAATTTTCGCTCTGACTCAAGGACTGCA 3660 Db
3613 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAAGAGTGAATTTGGAGTATCAGCTG 3672 QY
3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAAGAGTGAATTTGGAGTATCAGCTG 3720 Db
3673 GAAACATTTAGGTTCTTATTTCTAATGAAAAGTGAATAATGGAAGGCACTATTTCTCAA 3732 QY
3721 GAAACATTTAGGTTCTTATTTCTAATGAAAAGTGAATAATGGAAGGCACTATTTCTCAA 3780 Db
3733 CAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAAAAAG 3792 QY
3781 CAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAAAAAG -- 3838 Db
3793 GGTTTATTTAGTCGACGAAAAGAGAGCCCTGCTTTTACCCACACAGGTTCTCTCAGTAC 3852 QY
3839 -----AGTTCTCTCAGTAC 3855 Db
3853 AATGAGCTGAGCTGGCCTGGAGAAAGAGAAAGCTGCTGTGCAGAGCTAGAGAGAGCC 3912 QY
3856 AATGAGCTGAGCTGGCCTGGAGAAAGAGAAAGCTGCTGTGCAGAGCTAGAGAGAGCC 3915 Db
3913 CTTCAGAAAGACCCCATCGAGCTCCGGTCCGCCCGGAGGAAGCTGCCACCGCAAAAGCA 3972 QY
3916 CTTCAGAAAGACCCCATCGAGCTCCGGTCCGCCCGGAGGAAGCTGCCACCGCAAAAGCA 3975 Db
3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATTCGCGCC 4032 QY
3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATTCGCGCC 4035 Db
4033 ATCTGCGGTCGCGCAGAGCACCAGCCAGTGCATGAGCTGCTGGCCCGCCCATCCAGC 4092 QY
4036 ATCTGCGGTCGCGCAGAGCACCAGCCAGTGCATGAGCTGCTGGCCCGCCCATCCAGC 4095 Db
4093 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCCGGCTCTTAAAGGAACGATGCAC 4152 QY
4096 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCCGGCTCTTAAAGGAACGATGCAC 4155 Db
4153 CACAATTTCTCACCAGTTCAAGTGAAGTGAACATGCGAGCCACAAAGTGTGCTGTG 4212 QY
4156 CACAATTTCTCACCAGTTCAAGTGAAGTGAACATGCGAGCCACAAAGTGTGCTGTG 4215 Db
4213 TGTCTGGATACCGTGCACTTTGGACGCCAGGCAATCCAAATGTCTCGAATGTCAAGTGTG 4272 QY
4216 TGTCTGGATACCGTGCACTTTGGACGCCAGGCAATCCAAATGTCTCGAATGTCAAGTGTG 4275 Db
4273 TGTCAACCCCAAGTCTCCAGTGTCTGCCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4332 QY
4276 TGTCAACCCCAAGTCTCCAGTGTCTGCCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4335 Db
4333 ACACATTTCAACCGAGGCTTCTGCCGTGACAAAATGAATCTCCCGAGTCTCCAGACCAAG 4392 QY
4336 ACACATTTCAACCGAGGCTTCTGCCGTGACAAAATGAATCTCCCGAGTCTCCAGACCAAG 4395 Db
4393 GAGCCACGACAGCTTGCACCTGGAAAGGTTGATGAAGTGGCCAGGAATAACAAACGA 4452 QY
4396 GAGCCACGACAGCTTGCACCTGGAAAGGTTGATGAAGTGGCCAGGAATAACAAACGA 4455 Db
4453 GGAAGCAAGGCTGGGACAGAGATCATTTGCTCTGGAGGATCAGAAAGTCTCATTTAT 4512 QY
4456 GGAAGCAAGGCTGGGACAGAGATCATTTGCTCTGGAGGATCAGAAAGTCTCATTTAT 4515 Db

QY 4513 GACAAATGAACCCAGAGAGAGCTGACAGAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4572
 DB |||||
 QY 4516 GACAAATGAACCCAGAGAGAGCTGACAGAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4575
 DB |||||
 QY 4573 GACGGGATCTATCTATTATGTTGCGTGGTGGTTCGAACTCGCAAAATACAGCCAAA 4632
 DB |||||
 QY 4576 GACGGGATCTATCTATTATGTTGCGTGGTGGTTCGAACTCGCAAAATACAGCCAAA 4635
 DB |||||
 QY 4633 GCA----- 4635
 DB |||||
 QY 4636 GCAGATGCCCATACATACCTGAAGATGGAATCTCACCCGCGACACACCTGCTGGCCCGGG 4695
 DB |||||
 QY 4636 ----- 4635
 DB |||||
 QY 4696 AGAACCTCTACTTGTAGCTCCAGCTTCCCTTGACAAACAGCGCTGGGTACCGGCTTA 4755
 DB |||||
 QY 4636 -----GAAAAAGCAGAACTGATGCTAAATG 4662
 DB |||||
 QY 4756 GAATCAGTTGTCGAGGTGGGAGAGTTTCTAGGAAAAAGCAGAACTGATGCTAAATG 4815
 DB |||||
 QY 4663 CTTGGAAGCTCCCTGCTGAAGCTGGAAGTGAATGACGCTCTAGACATGAAGTCCAGCTG 4722
 DB |||||
 QY 4816 CTTGGAAGCTCCCTGCTGAAGCTGGAAGTGAATGACGCTCTAGACATGAAGTCCAGCTG 4875
 DB |||||
 QY 4723 CCCTTCAGTGACAGAGTGTGTTGGTGGGCAACGAGGAGGGCTCTACGCCCTGAAATGTC 4782
 DB |||||
 QY 4876 CCCTTCAGTGACAGAGTGTGTTGGTGGGCAACGAGGAGGGCTCTACGCCCTGAAATGTC 4935
 DB |||||
 QY 4783 TTGAAAACTCCCTAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4842
 DB |||||
 QY 4936 TTGAAAACTCCCTAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995
 DB |||||
 QY 4843 AAGGACCTGAGAGAGTCTCATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGAC 4902
 DB |||||
 QY 4996 AAGGACCTGAGAGAGTCTCATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGAC 5055
 DB |||||
 QY 4903 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCCCTGCCAGCCGACATCTCA 4962
 DB |||||
 QY 5056 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCCCTGCCAGCCGACATCTCA 5115
 DB |||||
 QY 4963 CCCAAATTTTGAAGCTGTCAAGGCTGCCATCTTGGGGCAGGCAAGTTGAGAC 5022
 DB |||||
 QY 5116 CCCAAATTTTGAAGCTGTCAAGGCTGCCATCTTGGGGCAGGCAAGTTGAGAC 5175
 DB |||||
 QY 5023 GGGCTCTGCATCTGTGCAGCATGCCAGCAAAAGTGTCTTCCGCTACACGAAAC 5082
 DB |||||
 QY 5176 GGGCTCTGCATCTGTGCAGCATGCCAGCAAAAGTGTCTTCCGCTACACGAAAC 5235
 DB |||||
 QY 5083 CTGAGCAATCTGCTATCCGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5142
 DB |||||
 QY 5236 CTGAGCAATCTGCTATCCGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
 DB |||||
 QY 5143 TTCACCAATTTACAGTATCTCTTGAACCAATAAATTTCTAGAAATCGACATGAAGCAG 5202
 DB |||||
 QY 5296 TTCACCAATTTACAGTATCTCTTGAACCAATAAATTTCTAGAAATCGACATGAAGCAG 5355
 DB |||||
 QY 5203 TACAGCTCTGAGGAATCTCTGATAGAAATGACATCTCTTGGCACCTGCTGTGTGCTG 5262
 DB |||||
 QY 5356 TACAGCTCTGAGGAATCTCTGATAGAAATGACATCTCTTGGCACCTGCTGTGTGCTG 5415
 DB |||||
 QY 5263 GCCTCTTCCAAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGCAGGCGAGAG 5322
 DB |||||
 QY 5416 GCCTCTTCCAAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGCAGGCGAGAG 5475
 DB |||||
 QY 5323 GAGTACTTGTGTTTCCACGAATTTGGAGTCTTCTGGATCTTTACGAAAGCTAGC 5382
 DB |||||
 QY 5476 GAGTACTTGTGTTTCCACGAATTTGGAGTCTTCTGGATCTTTACGAAAGCTAGC 5535
 DB |||||
 QY 5383 CGCAGAGCATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGCTTACAGAGAACCTAT 5442
 DB |||||
 QY 5536 CGCAGAGCATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGCTTACAGAGAACCTAT 5595
 DB |||||
 QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGCACGCTCCTCAGCA 5502

DB 5596 CTGTTTGTGACCCCACTTCAACTCACTCGAGTAAATTTGAGATCCAGGACGCTCTCTCAGCA 5655
 QY 5503 GGGACCCCTGCCCCGAGGTACTCTGACATCCGAAACCCGCGCTACCTGGGCCCCCTGCCATT 5562
 DB 5656 GGGACCCCTGCCCCGAGGTACTCTGACATCCGAAACCCGCGCTACCTGGGCCCCCTGCCATT 5715
 QY 5563 TCCTCAGAGAGCATTTACTTTGGGCTCTCATACAGGATAAATTTAAGGGTCATTGCTGC 5622
 DB 5716 TCCTCAGAGAGCATTTACTTTGGGCTCTCATACAGGATAAATTTAAGGGTCATTGCTGC 5775
 QY 5623 AAGGAAACCTCTGTGAAGGAGTCCGGCACTGAACACACCGGGCCGCTCCACCTCCCGC 5682
 DB 5776 AAGGAAACCTCTGTGAAGGAGTCCGGCACTGAACACACCGGGCCGCTCCACCTCCCGC 5835
 QY 5693 AG 5684
 DB 5836 AG 5837
 RESULT 13
 AF086824
 LOCUS
 DEFINITION Mus musculus rho/rac-interacting citron kinase (Crik) mRNA,
 complete cds.
 ACCESSION AF086824
 VERSION AF086824.1 GI:3599508
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 6954)
 AUTHORS Di Cunto,F., Calautti,E., Hsiao,J., Ong,L., Topley,G., Turco,E. and
 Dotto,G.P.
 TITLE Citron rho-interacting kinase, a novel tissue-specific ser/thr
 kinase encompassing the Rho-Rac-binding protein Citron
 J. Biol. Chem. 273 (45), 29706-29711 (1998)
 MEDLINE 99009084
 PUBMED 9792683
 REFERENCE 2 (bases 1 to 6954)
 AUTHORS Di Cunto,F., Calautti,E., Hsiao,J., Ong,L., Topley,G., Turco,E. and
 Dotto,G.P.
 TITLE Direct Submission
 JOURNAL Submitted (25-AUG-1998) Biology, University of Torino, Via Santena
 5 bis, Torino, TO 10126, Italy
 FEATURES
 Location/Qualifiers
 1..6954
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="spleen"
 1..6954
 /gene="Crik"
 772..6939
 /gene="Crik"
 /function="protein kinase"
 /note="Crik; rho/rac binding protein; alternatively
 spliced form of citron encoded by GenBank Accession Number
 AF039218; similar to Mus musculus citron-K encoded by
 GenBank Accession Number AF070066; similar to Rattus
 norvegicus citron-K encoded by GenBank Accession Number
 AF070065"
 /codon_start=1
 /product="rho/rac-interacting citron kinase"
 /protein_id="AAC72823.1"
 /db_xref="GI:3599508"
 /translation="MLKFKYGVNPPPEASASEPIASRSLNLFQKPLMTQQQMS
 ALSRGLMDALFALFECISQPALMKHVSFVKYSDTIAELRELPQARDPEVRSI
 VGCHFAEVOVREKATGVDYAMKMKKALLAQEQVSPFEERNILSRSTSPWIPOL
 QYAFQDKNLYLWMEYQGGDFLSILNRVEDQDESMIQFLAELILAVHSVHMGYV
 HRIKPNILIDRTGEIKLVFGSAAKNSKNKYDAKLPIGTDPDYMAPEVLTVNEDPR
 GTYGLDCDMSVGVVMEYVGYKTPFTFEGTSARTFNNIMNFQRFKFPDFDKVSSSELL

DLIQSLLCVQKRLKFEGLCPHFPPARTDWNIRNSPPFPVPTLKSDDDTNDFNPEPK
NSWAFILCVPAFLPSGSELPVGFYSKALGYLGRSESVSSVSDSPAKVSMKEL
LTKSELODSQDKHMEQEMTRLHRRVSEVAVLISOXELVKASETORSILLEODLAT
YITESSLSKSLEQARMEVSQEDDKALQLHDIRQSRKLEKEIQEYQAOVEMBLM
MNQLEDLVSARRSDLTSESDKSLRAAEFRKANECQHKMKAKQKPEVGEYS
KLEKINAEOLKIQBELQKLEKRAKSTEARTELLQNIROAKERAERELEKLUHREDSS
EGIKKIVBAEERLHKNKVRKTEEMRENNRDKDITQKSEIQQWADKILLEEK
HREAQVSAQHLNHLKQBOHYEERI KVLNDQIKKDLADKESLENMQRHEEAEHKG
KILSBOKAMINAMDKIRSLBOIRIVELSEANKLAANSILFTORNKMAOEMI SELRQ
KYLTOAGKLAQNRKLEBOLEKISHODSHKSKSLLETRLRVSLHEBEOKLELQ
RYLTLOLSQRESQLTALOARAALESQALQAKTELEETTAZAEERIQALTAHRE
IORKFDALNSCTVITDIEQUNQUTEDNAELNNQNFYLSQKDEASGNDEIVOLRS
SVHLUREITEREMOLTQOKTMEALKTTCTNMLEQVLDLEALNDELLEKQEWAMR
VYLGKSEOFECRVELQKMLDTEKQSRARADQRTESQVVELVAEKHKAIEIALQ
ALKEQKLAESLNDLLEKHAMLENVARSLOQLETERELKORLEEQEAKLQOQ
DIQKXNIFRLTOGLQBALDRALLKTERSDLEYQLENTQVLYSHKVMQETISQOTK
LIDFQAQMDQPAKKVQLOVYNELKLALEKAKCAELBEALOKTRTELRSABEA
HRKATDHPSPATARQOIANSAIVRSPHOPSAISLLAPSSRRKSSSPBESRR
LKERMHNNIHPFRNGLNKRATCAVCLDITVHFQKASCKLEQVCHPKKSTCLUPAT
CGLPAYIHTFEATCRDNMSPLQSKESPSSHLLEGVMKVPNRKNTAQQWDRKYI
VLEGKVLITDYNDAREAGORPVEFELCLPDGDVSIHGAVASELANAKADVPYILK
MESHPHTTCTWPGRTLYLLASPDPQKRWTALESVVAGRVSRKAEADAKLGNLIL
KLEGGDRLDNMCCTLPSDOVLVGTGEEGLYALNVLNLSLTHIPG:GAVFOIYIKDLE
KLIMLAGERALCLVDVKVKOSLAQSHLPAQPDVSPNIFPAVKGCHLFAAGKLENSL
CICAMPSPKVLIRNDNLSKICIRKEIETSPSCIHFTNYSILITGNKPYEDMKQ
RLTDLFLDKNHSRAPAVPASSNSFPVSVIVAGAGORBEYLLCFHFGEVFDVSYGR
RERTDOLKWSRULPAPVAREPYLFVTHFNSELEVIQIARSLGSFARAYLEIPNRYL
GPAISGAIYUPLASSQDKLRLVICCKGNLVKESQTEQHRVPSRSTRSPNRPPTNRYH
ITKRVASSPAPPEPSHPREPSTPHRYDRDREGLRLRDKSPGRPLEREKSPKRLMST
RRERSPGRLFEDSSRGRUPAGAVRTPLSQVKNVQDQSSV"

ORIGIN

Query Match		75.1%;	Score 4627;	DB 10;	Length 6954;
Best Local Similarity		85.7%;	Pred. No. 0;		
Matches 5332;		Conservative 0;	Mismatches 635;	Indels 254;	Gaps 7;
QY	1	ATGTTGAAGTCAAAATATGAGCGCGGGAATCCTTTGGATGCTGGTCTGCTCAACCCCAAT	60		
DB	772	ATGTTGAAGTTCAGTATGGTGTGCGGAACCCGCGGAGCGCAGTGCCTCCGAGCCCAAT	831		
QY	61	GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTTCCAGGGGAAAACCCCTTTATGACTCAA	120		
DB	832	GCCAGTCGGGCTCCAGGCTAAATCTCTTCTCCAGGGGAAAACCCGCTCATGACTCAA	891		
QY	121	CAGCAGATCTCTCTTTCCGAGAGAGGATATAGATGCCCTCTTTGTTCTCTTTGAA	180		
DB	892	CAGCAGATCTCTCTCTTTCCGAGAGAGGATGCTAGACGCCCTCTCTGCTCTCTTTGAA	951		
QY	181	GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT	240		
DB	952	GAGTGCAGCAACCCGCTGATGAGATGAAGCAGCTGAGCAGCTTTGTCAGAGATAT	1011		
QY	241	TCCGACACCATAGCTGAGTTAAGGAGCTCCAGCCTTCGGCAAGAGCTTCGAACTCAGA	300		
DB	1012	TCCGACACCATAGCCGAGTTGGGGAGCTGCAGCCTCGCGGAGAGACTTCGAAAGTTGGA	1071		
QY	301	AGTCCTTGTAGTTGTGCTCACTTTGCTGAAGTGCAGTGTGTAAGAGAGAAACCAACCCGG	360		
DB	1072	AGCCTTGTGGCTGTGCTCACTTCGCTGAAGTGCAGTGTGTTAGAGAGAGCGCACCCGG	1131		
QY	361	GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTATTGCGCCAGGAGCAGGTTTCA	420		
DB	1132	GACGTCATGCTATGAATCATGAAGAAGAGGCTTTGCTGCGCCAGGAAACAGGTTTCA	1191		
QY	421	TTTTTTGAGAGAGCGGACATATTTATCTGAGGACAGAGCCGCTGGATCCCCAATTA	480		
DB	1192	TTTTTTCGAGAGAGAGGAGCAATATTTATCTCGAGACAGAGCCGCTGGATCCCCAGTTA	1251		
QY	481	CAGTATGCTCTTCAGGACAAAATACCTTTTATCTGATGAGGAAATATCAGCCTCGAGGG	540		
DB	1252	CAGTATGCTCTTCAGGACAAAATACCTTTTATCTGATGAGGAAATATCAGCCTCGAGGG	1311		
QY	541	GACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATTAGATAAACCCTGATACAGTTT	600		

DB	1312	GATTTCTGCTGCTTCTTGAACAGATACGAGCAACCAATTAGATGAGATGATCAGTAT	1371		
QY	501	TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTCATCTGATGGATGATGCTGATCGA	660		
DB	1372	TACCTTGTGAGCTGATTTTGGCTGTTTCAAGCGTTCATCAGCGTTCACAGATGGGATATG	1431		
QY	661	GACATCAAGCCTCAGAACATTTCTGTTGACCGCAGCAGACACATCAAGCTGGTGGATTTT	720		
DB	1432	GACATCAAGCCCGAGAACATCTCTCATCGACCGGAGGAGATCAAGCTGGTGGATTTT	1491		
QY	721	GGATCTGCGCGGAAAATGAATTCACACAGATGGTGAATGCCAAATCCCGATTTGGGACC	780		
DB	1492	GGATCAGCGCTAAGATGAATTCACAAATAA---GGTGGATCCAAATCCCGATTTGGGACC	1548		
QY	781	CCAGATTACATGGCTCTCTGAAGTCTGACCTGTGTATGAACGGGATGAAAAGGACACCTAC	840		
DB	1549	CCGGAATTACATGGCTCCGGAAGTGTTCACCGTGTATGAACGAGGACCGAAGGGGCACATAC	1608		
QY	841	GGCTGGAATCTGATCTGGTGGTCTAGTGGGCGTGAATCCCTATGAGATGATTTATGGGAGA	900		
DB	1609	GGCTTGGACTGTGACTGGTGGTCTGTCTCGAGTGTGTTCCTTATGAGATGGTGTATGGGAAG	1668		
QY	901	TCCCTCTCCAGAGGGAACCTCTGCGCAGAACCTTCAATAACATTAATGAATTTCCACGCGG	960		
DB	1669	ACCCATTCACAGAGGGAACCTCCGCGGACCTTCAACAACTATGAACATTCAGACGGG	1728		
QY	961	TTTTTGAATTTCCAGATGACCCCAAAAGTGAGCAGTCACTTTCTTGATCTGATTCAAAGC	1020		
DB	1729	TTTTTGAATTTCCAGATGACCCCAAAAGTTAGCAGTGAAGTCTGATCTGCTTTCAGAGT	1788		
QY	1021	TTGTTGTGCGGCGAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCGCATCTTTCTTTC	1080		
DB	1789	CTGCTGTGTGTCAGAAAGAGAGACTGAAGTTCGAGGGTCTCTGCTGCCACCTTTCTTT	1848		
QY	1081	TCTAAAAATTGACTCGAAACAACTTCGTAACCTCTCTCTCCCTCCCTTCGTTCCACCCCTCAAG	1140		
DB	1849	GCCAGACGGGACTGGAACACACATCCGTAACCTCTCTCTCCCTTCGTTCCACCCCTCAAG	1908		
QY	1141	TCTGACATGACACTCCAATTTTGTATGAACCCAGAGAGAAATTCGTGGG--TTTCATCCTC	1199		
DB	1909	TCTGACATGACACTCCAATTTTGTATGAACCCAGAGAGAAATTCGTGGGCTTTTCATCCTC	1968		
QY	1200	TCCGTGCGAGTGAAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTCTGGGGTTTC	1259		
DB	1969	TGTGTGCGAGTGAAGCCCTC--GGTTTCTCAGGCGAAGAGCTGCCGTTTGTGGATTTTC	2027		
QY	1260	GTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTTGTGTCGGGTCTGGAATC	1319		
DB	2028	GTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTTGTGTCGGGTCTGGAATC	2087		
QY	1320	CCCTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAGA	1379		
DB	2088	CCCTGCCAAGGTTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAGA	2147		
QY	1380	CTCTCAGCAAGTGTCAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGT	1439		
DB	2148	CTCCAGGACAGTGTCAAGATGGAGCAGGAAATGACCCGGTTACATCGCAGAGTGT	2207		
QY	1440	AGAGTGGAGGCTGTGCTTGTAGTCAGAGAGGTGGAGCTGAAGGCTCTCAGACTCAGAG	1499		
DB	2208	AGAGTGGAGGCTGTGCTTGTAGTCAGAGAGGTGGAGCTGAAGGCTCTCAGACTCAGAG	2267		
QY	1500	ATCCCTCTCTGAGCAGGACCTTGTCTACATCACAAGATGCAAGTACCTTAAAGCGAAG	1559		
DB	2268	ATCCCTCTCTGAGCAGGACCTTGTCTACATCACAAGATGCAAGTACCTTAAAGCGAAG	2327		
QY	1560	TTTTGGACAAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTTCCA	1619		
DB	2328	TTTTGGACAAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTTCCA	2387		
QY	1620	TGATATCAGAGCAGAGCCGGAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCA	1679		

Db 6723 CTACCGAGACAGAGAGGGTCGACAGAGCTGCGCAGGAGCAAGTCTCCAGGCGCCCTCT 6782
 QY 5856 GGAGCGAGAGAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGTCCCGCGGAG 5915
 Db 6783 GGAGCGGAGAGATCCCGAGCGGATGCTCAGCAGCGGAGAGCGTCCCGAGGAG 6842
 QY 5916 GCTGTTTGAAGCAGCAGCAGCGGCGGCTGCTCGGAGCGGTCAGGAGCCCCGCTGTC 5975
 Db 6843 ACTGTTTGAAGCAGCAGCAGCGGCGGCTGCTCGGAGCAGTGAAGACCCACTGTC 6902
 QY 5976 CAGGTTGACAGGAGAGGCGAGAGTGCCTCTCAAGTTT 6016
 Db 6903 CCAGGTTAACAAGGCTGGGACCAAGTCTCAGTATAAGTCT 6943

 RESULT 14
 LOCUS AF039218 5952 bp mRNA linear ROD 02-JUL-1998
 DEFINITION Rattus norvegicus postsynaptic density protein (citron) mRNA,
 complete cds.
 ACCESSION AF039218
 VERSION AF039218.1 GI:2745839
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 5952)
 ZHANG, W., APPERSON, M.L., VASQUEZ, L.E. and KENNEDY, M.B.
 CITRON, a PSD-95-binding protein at glutamatergic synapses on
 inhibitory neurons
 Unpublished
 REFERENCE 2 (bases 1 to 5952)
 ZHANG, W., APPERSON, M.L. and KENNEDY, M.B.
 DIRECT SUBMISSION
 Submitted (18-DEC-1997) Biology, Caltech, Pasadena, CA 91125, USA
 FEATURES
 Location/Qualifiers
 1..5952
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..5952
 /gene="citron"
 613..5469
 /gene="citron"
 /function="interacting with Rho family GTPases"
 /codon_start=1
 /product="postsynaptic density protein"
 /protein_id="AAC25483.1"
 /db_xref="GI:2745840"
 translation="MFEQMARLHRRVSEVAVLSQKVELKASRTQSLLEODLATYI
 TECSLRSLERQNVSEDDKALQLLDHIREQSKLQEIKEQYQYQVEMRLMN
 QLEEDLVARRSDLYESELRESRLAAEFKRNCEQHKMKAKDLGKPNVEGSESL
 EKINAEQOLKIQLQELKAVKATEELLNTRQAKERABRELEKHNDEDSSE
 IKKVLVABERHSLKENVKLEMERENRLKDDITQKSEIQIMADKIIELEKHI
 EAQVQSLHEVLKQKQYVEEKIKVLDNQIKDLADKESLETMRQHEEAEHKGKI
 LSPQKAMINADSKISLSQRIVELSEANKLAANSILFTORNKAQEMISELROKF
 YLETOAGKLEAONKLELELEKTSKSHODHDKNRLLETRLEVSLEHTEKLEKRO
 LTELQSLQERESQLTALAQAALLESQRLQAKTELEETAEAEIEIQAHTAREIQ
 RKFDLARNSCTVITDLEQNLQNTDNLNQNFIYSKQDLDEPAGANDEIVQLRSEV
 DLHREITREMTVLTSQKTMELKTTCTMLEEQVMDLEALNDELLEKQWAEARV
 LGDEKQFQRELQMDLDEKQSPADADITRESQVVELAVKEHAEILALQOAL
 KEQKLXESLSDKRLKHAMLENARSLQKLETERELKQRLLEQALQOQMDL
 QKNIHRLTQGLLEADLLKTERSDLEYOLENTQVLYSHEKVMGEGTISQOTKLI
 DFLQAKMDPAKGLSRKEDPALTPQVLOYNELKALEKELKARCALEELAKOK
 TRLELSAEAEAAHRAKTHDPHSTPATRQIKSAIVRSEPHQSPMSLILAPSSR
 KEASPTPEFSLKERNMHNIHPRFNVLGNMRAKCAVCLDTVHFGRQSKLSCQV
 MCHPKSCITLPATCGLPAYATHTEAFCRDKVSSPGLQSKPSSSLHLEGMWKVFRN
 NKRGQQQWDRKYTVLEGSKVLIVDNEAREAGRPVEEFELCLPDGDSVTHGAVASEL
 ANTKADVPYILKWEHSHPTTCWPGRTLYLLAPSPFDKQWVTALESVVAVGVRSEK
 AEADAADCTSCERLPVWVEKLLNSLLKLEGGDRLDMNCTLPFSDQVVLVTERGLY
 ALNVLKNSLTHIFGIVAGFYIYIKLEKLLMIAGERALCLVDVKVKQSLAQSHLP

AQPDSNPNIPEAVKGCGLPFAAGKIEIENSLCICAMPKSVILRYNDNLSEKTRKELET
 SEFCSTHFNYSILIGTNKFYBIDMKQYTLDFLNDKNDHSLAPAVPASTSTSPFVSI
 VQANSTQORBEYLLCFHEFGVFDVDSRRSRDLDLKWRLPLAFAPVASTSTSPFVSI
 LEVETQARSLGTSPARVLEIENPYRLGPAISSGAIYIASSYQDLKRVICCKGNLVK
 ESGTEOHRVPSSTSPKRPPTYNBHIKRVASSPAPPEPSHPPESTPHRYRRE
 GRTELPRDKSPGRLEREXSPGRMLSTRERSPGRLFEDSRGRLPAGAVRTPLSQVN
 KWDQSSV
 misc_feature 3052..3096
 /gene="citron"
 /note="may be excised by alternative splicing"
 misc_feature 4066..4110
 /gene="citron"
 /note="may be excised by alternative splicing"
 ORIGIN
 Query Match 58.5%; Score 3602.6; DB 10; Length 5952;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 4171; Conservative 0; Mismatches 444; Indels 252; Gaps 5;
 QY 1399 AAGATGAGCAGGAGAAATGACCGGTTACATCGGAGAGTGTACAGGTGGAGGCTGTGCTT 1458
 Db 610 ATGATGGAGCAGGAAATGGCCCGTTGTCATCGCAGAGTGTACAGGTGGAGGCTGTGCTT 669
 QY 1459 AGTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGATCCCTCTGGAGCAGGAC 1518
 Db 670 AGTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGATCCCTCTGGAGCAGGAC 729
 QY 1519 CTTGCTACCTACATCAGAGATGCGTAGCTTAAAGCGAGTTTGGAGCAGCAGCAGATG 1578
 Db 730 CTTGCCACCTACATCAGAGATGCGTAGCTTAAAGCGAGTTTGGAGCAGCAGCAGATG 789
 QY 1579 GAGGTGTCCAGAGAGGATGACAAAGCAGCTCAGGCTTCTCCATGATATCAGAGAGCAGAGC 1638
 Db 790 GAGGTGTCTCAGGAGGATGACAAAGCAGCTTCTCCATGATATCAGAGAGCAGAGC 849
 QY 1639 CGGAAGTCTCAAGAAATCAAGAGCAGGAGTACAGGCTCAAGTGGAGAGAAATGAGTTG 1698
 Db 850 CGGAAGTCTCCAGGAGATCAAGGAGCAGGAGTACAGGCTCAGGTGGAGGAGATGAGGCTG 909
 QY 1699 ATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAGAAAGACGAGTGTATCTACGAA 1758
 Db 910 ATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAGAAAGACGAGTGTATCTACGAG 969
 QY 1759 TCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGATTAAGCGGAAAGCGCAGAGATGT 1818
 Db 970 TCTGAGCTGAGGAGTCTCGGCTCGCGCCGAGGAATTCAGCGGAAAGCGCAGAGATGT 1029
 QY 1819 CAGCATAAATCTTGAAGGCTTAGGATCAAGGGAAGCTCAAGTGGAGAGATATGCGAAA 1878
 Db 1030 CAGCATAAATCTTGAAGGCTTAGGATCAAGGGAAGCTCAAGTGGAGAGATATGCGAAA 1089
 QY 1879 CTGAGAGAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAGTCCAAGAGAACTCGAG 1938
 Db 1090 CTGAGAGAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAGTCCAAGAGAACTCGAG 1149
 QY 1939 AAGGCT-----GCAAG 1950
 Db 1150 AAGGCTGTAAAGCCACACCGAGGCCACAGAGCTCTGAGAGAACTCCGCGAGGCAAG 1209
 QY 1951 GAGCAGCCGAGAGGAGCTGGAGAACTGCGAGAACTGCGAGAGAGATTTCTTGAAGGATC 2010
 Db 1210 GAGCAGCTGAGAGGAGCTGGAGAACTGCGAGAACTGCGAGAGAGATTTCTTGAAGGATC 1969
 QY 2011 AGAAAGAGCTGGTGGAGAGTGGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAG 2070
 Db 1270 AAAAAAGAACTGGTGGAGAGTGGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAG 1329
 QY 2071 CTAGAGCCATGGAGCTGAGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAG 2130
 Db 1330 CTAGAGCCATGGAGCTGAGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAG 1389
 QY 2131 CAGATCCAGCAGATGGCTGTATAAATTCAGGAGCTCGAAGAGAACTCGGAGGCCCAA 2190

Db 1390 CAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTGGAGGAAGCATCGGAGGCTCAG 1449
Qy 2191 GTCTAGCCGACACCTTAGAGTGCACCTGAAACAGAAAGAGCAGCAGCTATGAGGAAAG 2250
Db 1450 GTCTCAGCTCAACACTTAGAGTACACCTGAAAGCAGAAAGGACAGCAGCTATGAGGAAAG 1509
Qy 2251 ATTAAGTGTTCGAACAATCAGATAAGAAAGACCTGGCTGACAGAGGACACTGAGAAC 2310
Db 1510 ATCAAAGTGTTCGAACAATCAGATAAGAAAGGACTTGGCCGACAAAGGAGGCTAGAGACC 1569
Qy 2311 ATGATGCGAGACACGAGGAGGAGGCCATGAGAGGAGCCAAATTTCTCAGCGACAGAG 2370
Db 1570 ATGATGCGAGGACACGAGAGGAGGCCACGAGAAGGCGCAGATCCTCAGCGAGCAGAG 1629
Qy 2371 GCGATGATCAATGTATGGATTCCAGATTCAGATCCCTGGAAACAGAGGATTGTGAACTG 2430
Db 1630 GCGATGATCAACGCAATGGATTCCAGATTCAGATCCCTGGAGCAGAGGATTGTGAGCTG 1689
Qy 2431 TCTGAAGCCAAATAAATTCGACCAATATGACAGCTCTTTTATCCCAAGGAACATGAGGCC 2490
Db 1690 TCAGAAGCCAAACAAGCTTTCGGCAACAGCAGCTCTTCACCGACGAGGAATGAGGCC 1749
Qy 2491 CAAGAAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGGCTGG 2550
Db 1750 CAGGAGAGATGATCTCAGAACTCAGGCGACAGAAATTTTACCTGGAGACACAGGCCGG 1809
Qy 2551 AAGTTGGAGGCCAGAACCGGAAATTCGGAGGAGCAGCTGGAGAAATCAGCCACCAAGAC 2610
Db 1810 AAGCTGGAGGCCAGAACCGGAAATTCGGAAGAGCAACTAGAGAAGATCAGCCACCAAGAT 1869
Qy 2611 CACAGTGACAAGATCGGCTGTGGAACTGGAGACAAGATTCGGGAGGTCAGTCTAGAC 2670
Db 1870 CACAGTGACAAGATCGGCTGTGGAGCTGGAGACAAGTTGAGAGAGGTCAAGCTGGAG 1929
Qy 2671 CACGAGGAGCAAACTGGAGCTCAAGCGCCAGCTCACAGAGCTACAGCTCTCCCTGCGAG 2730
Db 1930 CATGAGGAGCAAGCTGGAGCTGAAGCGGCAGCTCACGAGCTCAGCTGTCCCTCGAC 1989
Qy 2731 GAGCGGAGTCACAGTTGACACCTTCAGGCTGACGGGTGACGGGGCCCTGGAGAGCCAGCTT 2790
Db 1990 GAGCTGTAGTCCAGCTGACGCTTCAGGCGCCCTGAGGCGCCCGGGCTGCCCTTGAAGCCAGCTC 2049
Qy 2791 CGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGAGAGGAGCTGGAAGAGAGATCCAGGCA 2850
Db 2050 CGCCAGGCGAAGAGCTGAGGAGACAATGAGAGAGCGGAGAGAGATCCAGGCA 2109
Qy 2851 CTCACGGCACATAGAGTGAATTCAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACT 2910
Db 2110 CTCACGGCACATAGAGTGAATTCAGCGCAAAATTCGATGCCCTTCGCAACAGCTGTACT 2169
Qy 2911 GTAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAGGACACGCTGAATCTCAAC 2970
Db 2170 GTATCACCAGCTTGAGGAGCAGCTGAACAGCTCAACGAGGACCAACCGAGCTCAAC 2229
Qy 2971 AACCAAACTTCTACTGTTCAAAACAACTCGATGAGGCTTTCGGCGCCAAACGACAGATT 3030
Db 2230 AACCAAACTTCTACTGTTCAAAACAACTCGATGAGGCTTTCGGGTGCCAAACGATCAGATT 2289
Qy 3031 GTACAACTCGAAGTGAATGAGACCATCTCCCGCGGAGATCACGGAACGAGAGATGAG 3090
Db 2290 GTGCACTCGAAGCGAGGTGAGACCATCTCCCGCGTGATCACGAGCGGAGATGAG 2349
Qy 3091 CTTACAGCGCAGAAACGATGAGGCTCTGAAGACCACTGACACCATGACCATGCTGGAGAA 3150
Db 2350 CTCACAGCGCAGAAACGATGAGGCTCTGAAGACCACTGACACCATGCTGGAGAG 2409
Qy 3151 CAGGTCAATGGATTGGAGCCCTTAACGATGAGCTGCTAGAAACAGCGCGAGTGGAG 3210
Db 2410 CAGGTCAATGGACCTGGAGCCCTTGAACGACGAGCTGCTGGAGAGGAGCGCCAGTGGAG 2469
Qy 3211 GCTGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGTGCTGGGTTCGAGAGCTG 3270
Db 2470 GCTGAGGAGCGCTCTCGGGACGAGAAATCTCAGTTTGTGATGCTGAGGCTGAGAGCTA 2529

Qy 3271 CAGAGAAATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCCGATCAGCGGATCACCAG 3330
Db 2530 CAGAGGATGCTGGACACCGAGAAACAGAGCAGGGCTAGAGCTGATCAGCGGATCACCAG 2589
Qy 3331 TCTCGCAGGTGTGAGCTGGCAGTGAAGGACACAAGGCTGAGATTCTCGCTCTGCGAG 3390
Db 2590 TCTCGCAGGTGTGAGCTGGCAGTGAAGGACACAAGGCTTACAGCAGAAATCGGAG 2649
Qy 3391 CAGGCTCTCAAAAGACAGAAAGCTGAAGGCCGAGAGCTCTCTGACAAAGCTCAATGACCTG 3450
Db 2650 CAGGCTCTCAAAAGACAGAAAGCTGAAGGCTGAGAGCTCTCGGACAAAGCTCAAGACCTG 2709
Qy 3451 GAGAAGAGCATGCTATGCTTGAATGAATGCTCCGAAAGCTTACAGCAAGCTTGGAGACT 3510
Db 2710 GAGAAGAGCAGCCATGCTTGAATGAATGCTCCGAAAGCTTACAGCAGAAATCGGAGACA 2769
Qy 3511 GAACGAGCTCAACAGAGGCTTCTGGAAGACCAAGCCAAATTTACAGCAGCAGATGAGAC 3570
Db 2770 GAGCGGAGCTCAACAGAGGCTTCTGGAGAGCAAGCAGCTTGCAGCAGAGATGAGAC 2829
Qy 3571 CTGCAAGAAATCACAATTTTCCGCTGACTCAAGGACTGCAAGAGCTCTAGATGCGGCT 3630
Db 2830 CTGCAAGAAACACATCTTTCAGACTGACACAAGGCTGCAAGGCGCTGAGCCCGGGC 2889
Qy 3631 GATCTACTGAAGACAGAAAGTGTGAGTATCAGCTGGAACATTCAGGTTCTC 3690
Db 2890 GATCTGCTGAAGCCGAAAGGAGCGACTGGAATATCAGCTGGAACATTCAGGTTCTC 2949
Qy 3691 TATTCTCATGAAAGGTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCAATTGAT 3750
Db 2950 TACTCAACAGAGAGGTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCAATTGAT 3009
Qy 3751 TTTCTGCAAGCCAAATGGAACCTGCTGCTTAAAGAAAGGCTTATTATTAGTCGACGG 3810
Db 3010 TTTCTGCAAGCCAAATGGAACCTGCTGCTTAAAGAAAGGCTTATTATTAGTCGACGG 3069
Qy 3811 AAAGAGACCTCTTTTACCCACACAGGTTCTCTGCAGTACAAATGAGCTGAAAGCTGCC 3870
Db 3070 AAAGAGACCTCTGCTTTGCCACACAGGTTCTCTGCAGTACAAATGAGCTGAAAGCTGCC 3129
Qy 3871 CTGGAAGAGAGAAAGCTGCTGTGCAAGCTAGAGAAAGCTTTCAGAAAGACCCGATC 3930
Db 3130 CTGGAAGAGAGAAAGCTGCTGTGCGAGCTGGAAGAGCTTTCAGAAAGACCCGATC 3189
Qy 3931 GAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3990
Db 3190 GAACTCCGCTCCGCGGAGGAGCTGCTCACCGCAAGCCAGGACCCAGGACCCAGCACC 3249
Qy 3991 TCCACGCGCAGCCACCGGAGGAGCAGATCGCCATGTCGCCCATGTCGCCCATGTCGCCCATG 4050
Db 3250 TCTACGCGCAGCCTGCGAGGAGCAGATCGCCATGTCGCCCATGTCGCCCATGTCGCCCATG 3309
Qy 4051 CACGAGCCAGTGCATGAGCTGCTGGCCCGGCGCATCGAGCGCAGAAAGGAGTCTTCA 4110
Db 3310 CACGAGCCAGTGCATGAGCTGCTGGCCCGGCGCATCGAGCGCAGAAAGGAGGCTCA 3369
Qy 4111 ACTCCAGAGAAATTTAGTCGGGCTTTAAGGAAAGCAGATGCAACCAATATTCCTCAGCGA 4170
Db 3370 ACTCCAGAGAGTTACGCGGCTTTAAGAGAGCTATGCAACCAACATTCCTCAGCGG 3429
Qy 4171 TTCAACCTTAGAGCTGAACATGCGAGGACCAAAAGTGTGCTGTGTCTGAGATACGCTGAC 4230
Db 3430 TTTAACTGGGCTTGAATATGCGAGGCCACCAAGTGGCGCTGTGTCTGTGATCTGTGAC 3489
Qy 4231 TTTGGAGCCAGAGCATCCAAATGCTCGAATGTCAGGTGATGTCAACCAAGTGTCTC 4290
Db 3490 TTTGGAGCCAGGATCCAAATGCTCGAATGTCAGGTGATGTCCCATCCCAAGTGTCTC 3549
Qy 4291 ACCTGCTTGGCAGCAGCTCGCGCTTGCCTGCTGAATATGCCACACACTTTCACCGAGGCC 4350
Db 3550 ACCTGCTTGGCAGCAGCTCGCGCTTGCCTGCTGAATATGCCACACACTTTCACCGAGGCC 3609

QY 4351 TTCTGCGGTGACAAAATGAATGAACTCCAGAGTCTCCAGACCAAGAGAGCCAGAGAGAGCTTTG 4410
DB 3610 TTCTGCGGACACAAAGTGAGCTCCCGGGCTCCAGAGCAAGAGAGCCAGAGAGCTTTG 3669
QY 4411 CACCTGGAAGGGTGGATGAGGTGCGCAGGATTAACAAACGAGGACAGACAGCTGGGAC 4470
DB 3670 CACTGGAAGGGTGGATGAGGTGCGCAGGATTAACAAACGAGGACAGACAGCTGGGAC 3729
QY 4471 AGAAGTACATTTCTCTGGAGGATCAAAAGTCTCTCATTTATGACAAATGAAGCCAGAGAA 4530
DB 3730 AGAAGTACATTTCTCTAGAGGGTTCAAAAGTCTCTCATTTATGACAAACGAGCCAGAGAA 3789
QY 4531 GCTGGACAGAGGCGGTGGAGAAATTTGAGCTGTGCTTTCCGACGGGATGTATCTATT 4590
DB 3790 GCTGGACAGAGGCGGTGGAGAAATTTGAGCTGTGCTTTCCGACGGGATGTATCTATT 3849
QY 4591 CATGGTCCCTTGGTCTTCCGAACCTCGCAAAATACAGCCAAAGCA 4635
DB 3850 CATGGTCCCTTGGTCTTGAACCTGCAAAATACAGCCAAAGCAGATGCCATACATA 3909
QY 4636 ----- 4635
DB 3910 CTGAAGATGAGATCCCATCGCATACACCTGCTGGCTGGGAGAACCTCTACCTGCTG 3969
QY 4636 ----- 4635
DB 3970 GCACCCAGCTTCTCTACAAAGCAGCGCTGGGTACAGCCCTAGAGTCTGTCTGCGAGGT 4029
QY 4636 -----GAAAAAGCAGAGCTGAGCT 4656
DB 4030 GCGAGAGTTTCTAGGAAAGGAGCAGAGCGGATGCTGTCATGGGAGCTGACTTCTTGTGAG 4089
QY 4657 -----AACTGCTTGGAACTCCCTGCTGAACTGGAAGGTGAT 4695
DB 4090 CGTCTCCCGTGGGTGAGAAATTTACTTGAACCTCTCTGCTGAAGCTGGAAGGTGAC 4149
QY 4696 GACCGTCTAGACATGAATGCGCTGCGCTTCAAGTACAGAGTGGTGTGTTGGGCGACC 4755
DB 4150 GACCGCTTGAATGAGTACCTGCTGCGCTTCAAGTACAGAGTGGTGTGTTGGGCGACC 4209
QY 4756 GAGGAAGGCTCTACCGCTGAGTCTTGAAGAACTCCCTAACCCATGTCCTGAGGAAT 4815
DB 4210 GAGGAGGCTCTATGCGTGAACCTTGAAGAACTCCCTAACCCATGTCCTGAGAGTGAAGGTGAC 4269
QY 4816 GGAGCAGTCTTCCAAATTTATATTAATCAAGGACCTGGAGAGCTACTCATGATACAGGA 4875
DB 4270 GGAGCAGTCTTCCAGATTTACATCATCAAGAGCTGGAGAGCTGCTCATGATACAGGG 4329
QY 4876 GAAGAGCGGCTCTGCTGCTTGTGAGAGTGAAGAAAGTGAACAGTCTGCTGCGCCAGTCC 4935
DB 4330 GAAGAGCGGCTCTGCTGCTTGTGAGAGTGAAGAAAGTGAAGAGTCTGCTGCGCCAGTCA 4389
QY 4936 CAGCTGCTGCGCAGCGCCAGATCTCACCAATTTTGAAGCTGTCAAGGCTGCCAC 4995
DB 4390 CAGCTTCTGCGCAGCTGAGCTCTCCCCCAACATTTTGAAGCGTCAAGGCTGCCAC 4449
QY 4996 TTGTTTGGGCGGCAAGATGAGAAAGGCTCTGATCTGTGAGCCATGCTCCAGCAAA 5055
DB 4450 CTGTTTGTGCTGGCAAGATGAGAAAGGCTCTGATCTGTGCGCCATGCTCCAGCAAA 4509
QY 5056 GTGCTCATTTCTGGTACAAAGAACTCAGCAATCTGATCTGCGGAGAGATAGAG 5115
DB 4510 GTGCTCATTTCTGGTACAAAGAACTCAGCAATCTGATCTGCGGAGAGATAGAG 4569
QY 5116 ACTCAGAGCCCTGAGCTGTATCCACTTCAACAAATTAAGATCTCTCATTTGGAACCAAT 5175
DB 4570 ACTCTGAGCCCTGAGCTGTATCCACTTCAACAAATTAAGATCTCTCATTTGGAACCAAC 4629
QY 5176 AAATCTAGAAATGACATGAAGAGATGACGCTGAGGAATTTCTGATGAAGATGAC 5235
DB 4630 AAGTCTATGATGACATGAAGAGATGACGCTGAGGAATTTCTGAGCAAGATGAC 4689
QY 5236 CATTCTTGGCAGCTGCTGTTTGGCGCTCTTCCAAAGCTTCCCTGTCTCAATGCTG 5295

DB 4690 CATTCTTGGCAGCTGCTGTTTGGCATCTTCGACCAACAGCTTCCCGCTCTCCATCGT 4749
QY 5296 CAGGTGAACAGCGCAGAGGAGAGTACTTGTGTTTCCACGAATTTGGAGTG 5355
DB 4750 CAGCAAAACAGCAGCGGAGAGAGTACTTGTGTTTCCACGAATTTGGGAGTG 4809
QY 5356 TTCTGTGATTTTACGGAAGAGTGTAGCGGACAGATCTCAAGTGTGAGTGTGCTTACCT 5415
DB 4810 TTCTGTGATTTTATGGAAGAGTGTAGCGGACAGATCTCAAGTGTGAGTGTGCTTACCT 4869
QY 5416 TTGCGCTTTCCTACAGAGAACCTTCTGTTGTGACCCACATTTCACTCACTCGAAGTA 5475
DB 4870 CTGCGCTTTCCTACAGAGAACCTTCTGTTGTGACTCACTTTAACTCCCTCGAAGTA 4929
QY 5476 ATTGAGATCCAGGACAGCTCTCTCAGAGGAGCCCTGCGGAGCGTACCTTGGACATCCCG 5535
DB 4930 ATTGAGATCCAGGACAGCTTCTCTCACTGGGAGCCCTGCGGAGCGTATCTTGAATCCCA 4989
QY 5536 AACCGCGCTTACCTGGGCTTCTCTCAGAGCGATTTTACTTGGCGTCTCTCATAC 5595
DB 4990 AACCTCTGCTTACCTGGGCTTCTCTCAGAGCAATTTACCTGGCGTCTCTCATAC 5049
QY 5596 CAGGATAAATTAAGGGTCAATTTCTCAAGGAGAACTCTGTAAGAGTCTCGGCACTGAA 5655
DB 5050 CAGGACAGTTTAAAGGTCTATGCTGCAAGAGAACTCTGTAAGAGTCTCGGCACTGAG 5109
QY 5656 CACACCGGCGCTTCTCAGCTCCCGAGCGAGCCCAACAGCGAGGCGCCACCCAGCTAC 5715
DB 5110 CAGCAGCGCTGCGCTCCACATCCC---GCAGCGCTTAAACAGCGAGGCGCCACCAACATAC 5166
QY 5716 AACGAGCAGATCACCAAGCGCTGGCTCCAGCGCGCGCCGCGGAGGCGCCAGCGCAC 5775
DB 5167 AACGAGCAGATCACCAAGCGCTGGCTCCAGCGCGCGCCAGCGCGCCAGCGCAC 5226
QY 5776 CCGCAGAGCAAGCAGCACCCCTACCTACC-----GCGAGGCGGAGCCGAGCTGCGC 5829
DB 5227 CCGCAGAGCAAGCAGCACCCCTACCTACCGAGACAGAGGCGCGGACAGAGCTGCGC 5286
QY 5830 AGGACAGCTCTCTGCGCGCGCTGGAGCAGAGAGTCTCCCGCGCGGAGTCTCAGC 5889
DB 5287 AGGACAGCTCTCAGCGCGCTCTCTGAGCGGAGAGTCTCAGCGCGGAGTCTCAGC 5346
QY 5890 ACGCGAGAGAGGCGTCCCGCGGAGGCTGTTTGAAGACAGAGCAGGCGCGGCTGCT 5949
DB 5347 ACGAGAGAGAGCGTCCCGCGGAGGAGTCTTGAAGACAGCAGCAGGCGCGGCTGCT 5406
QY 5950 GCGGAGCGCTGAGGAGCGCTGCTCCAGGTTAAACAGGAGAGGCGGAGGCTGCT 6009
DB 5407 GCGGAGCGCTGAGGAGCGCTGCTCCAGGTTAAACAGGTTAAACAGGTTGAGGACAGTCTCAGTA 5466
QY 6010 CAAAGTTT 6016
DB 5467 TAAGTCT 5473

RESULT 15
AF070066

LOCUS Mus musculus Citron-K kinase mRNA, partial cds. 4967 bp mRNA linear ROD 31-JUL-1998
DEFINITION Mus musculus Citron-K kinase mRNA, partial cds.
ACCESSION AF070066
VERSION AF070066.1 GI:3360513

KEYWORDS
SOURCE

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 4967)
Madaule,F., Furuyashiki,T., Reid,T., Ishizaki,T., Watanabe,G.,
Morii,N. and Narumiya,S.

TITLE
JOURNAL

A novel partner for the GTP-bound forms of rho and rac
FEBS Lett. 377 (2), 243-248 (1995)
MEDLINE 96128238

8543060
2 (bases 1 to 4967)
REFERENCE
AUTHORS
Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuoka, T., Bito, H.,
Ishizaki, T., and Narumiya, S.
TITLE
Role of Citron kinase as a target of the small GTPase Rho in
cytokinesis
JOURNAL
Nature (1998) In press
3 (bases 1 to 4967)
REFERENCE
AUTHORS
Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuoka, T., Bito, H.,
Ishizaki, T., and Narumiya, S.
TITLE
Direct Submision
JOURNAL
Submitted (04-JUN-1998) Pharmacology, Faculty of Medicine, Kyoto
University, Yoshida-Konoe, Sakyo-ku, Kyoto 606 8315, Japan
FEATURES
Location/Qualifiers
1..4967
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
41..4928
/function="binds to Rho and regulates cytokinesis"
/note="kinase variant form of Citron; approximately the
first 50 amino acids belong to the kinase domain;
alternatively spliced form of Citron, encoded by GenBank
Accession Number AF039218; similar to Rattus norvegicus
Citron-K encoded by GenBank Accession Number AF070065"
/codon_start=3
/product="Citron-K kinase"
/protein_id="AAC27933.1"
/db_xref="GI:3360514"
/translation="PFVPTLKSDDDTSNFDPEKNWSVSSVCSLSPSGSEBELPFV
GFSYKALGYLRSVSSVSLDSPAKVSSMEKLLIKSELQDSQDKHMEQEMTLR
HRRVEVALVSKVELKASETQSLLEQDLATVITCSLSLKRSLEQARMEVQEDD
KALQLHDIRESRLKQIKQBYQAOVEVRLMNMQLLEBLVSARRSDLYSELRE
SLAAEFKRNQIRAKAEKQKPEVGEYSKLEKINAEQQLKLEQLKLEKAV
KASTATELLQNRQAKAEKLEKHNDESEGIKKLVEAELEBEKHEKQVSA
QHEVLKQKEQHEKIKVLDNQIKDLADKESLENNMORHEBEKHKILSEOKA
MINADSKTSLRSQRIVELSEANKLAANSII.FTORNMAQEMISELQKQFYLETQA
KLEAQNKLRLQLEKIKI SHQSDKSLLELRLREVSLEHEQKLEKQLTLEQL
SQLESQTLQARAALLESQLRQAKTELETTAEABEEIQALTFAHDEIQKEDAL
RNSCTVITLLEQLNQLTEDNALNNQNFYLSKQDDEASGANDEIVQLRSEVDHLRE
ITEREMQTSQKQMEALKTQTMLEEQVLDLEALNDELLEKEROWEAWRSVLGDEKS
QPEVRSLQMLDTEKQSRADQITQESQVVELAVKEHAEILALQKALKEQKLS
ABSLQDLNDELKHAMLEMNARSQKLEITRELKQRLLEEQKALQOQMDLQKNHIF
RLTQGLQALDRDLKTERSDLEOLENIQVLYSHEKVMEGTISQOQTKLIDFLQAK
MDQAKKVPLOYNELKALEKAKAELEALQKTRIELRSARESEALHKKADHP
HESFATRAQITAMSAIVRSPHQPMSALLAPPSRRKESSTPEFSRLKERVHN
IPIHENVGLNNRKAQVCLDITVHFGQASKRCLEQVCHPEKSCSTCLPATCGLPAEYA
THFTAFCDKYNPSGLQSKPEGSLHLEGWVKVPRNNKQGGQMDRYIIVLEGSKVL
IYDNEARAGQRPVEFEELCLPDGDVSIHGAVGASELANTAKADVPYILRMESHPTT
CWPGRTLVLAPSPFDKORWTVLSSVAVGVRVREKAEADAKLGNLSLLKLEGDRL
DMNCTLPSPDOVVLVGTBGLVALNVLKNSLTHIPGIGAFQIYI.IKOLEKILMAGE
ERALCLVDVKVKQSLAQSHI PAQPDVSPNIFPAVKGCHLFAAGKIENSLCICAMPS
KVYILRYNDNLSKICIRKEIETSEFPCSIHFTNYSILIGTNKFPYFVDSKRSRTDLD
KNDHSLAPAVFSSNSFVSIQANSAGQREYLLCFHFEFGVFDVTSKRSRTDLDK
WRLPLAFAYEPYLVFTHNLSLVIEIQARSSLSGPARAYLEIPNRYLFGAITSIGA
IYLAASYQDKLRYICCKGNLVKESGTEQHRVSTSRSPKRPGETNYNEHTTKRVASS
PAPPSPHSPRPSTPHRYRDRGRTLELRDKSPGRPLEREKSPGRMLSTRERSPGR
LFEDSSRGRPLPAGAVRTPLSQNVKWDQSSV"

Query Match 58.4%; Score 3593.8; DB 10; Length 4967;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 4249; Conservative 0; Mismatches 527; Indels 279; Gaps 6;
QY 1118 CCCCCTTCGTTCCACCCCTCAAGCTGACGATGACACCTCCAAATTTTGATGAACACAGAGA 1177
DB 1 CCCCCTTCGTTCCACCCCTCAAGCTGACGATGACACCTCCAAATTTTGATGAACACAGAGA 60
QY 1178 AGAATTCGTGGGTTTCATCCTCTCCGTGCCAGCTGAGCCCTTCAGGCTTCTCGGGTGAAG 1237
DB 61 AGAATTCGTGGGTTTCATCCTCTCTGTGTGCCAGCTGAGCCCTTCAGGCGGAAG 120

ORIGIN

1238 AACTGCCGTTTGTGGGTTTTCGTACAGCAAGCAGCTGGGATTCTTGGTAGATCTGAGT 1297
121 AGTGCCTGTTGGGATTTTCGTACAGCAAGCAGCTGGGTTAATCTTGGTAGATCTGAGT 180
1298 CTGTGTGTGGGTTCTGGACTCCCTCCCAAGACTAGCTCCATCGAAAGAAACTTCTCA 1357
181 CTGTGTGTGGGTTCTGGACTCCCTCCCAAGACTAGCTCCATCGAAAGAAACTTCTCA 240
1358 TCAAAAGCAAGAGCTTACAAGACTCTCAGGACAAGTGTACAGAGTGGAGCGAAGTGA 1417
241 TCAAAAGCAAGAGCTTCCAAGACTCCAGGACAAGTGTCAAGAGTGGAGCGAAGTGA 300
1418 CCGGTTTACATCGAGAGTGTACAGAGTGTGAGGCTGTCTTGTAGTCAGAGAGGAGTGA 1477
301 CCGGTTTACATCGAGAGTGTACAGAGTGTGAGGCTGTCTTGTAGTCAGAGAGGAGTGA 360
1478 TGAAGGCTCTGAGACTCAGAGATCCCTCTCGAGAGAGGAGCTTGTCTACCTACATCAG 1537
361 TGAAGGCTCTGAGACTCAGAGATCCCTCTCGAGAGAGGAGCTTGTCTACCTACATCAG 420
1538 AATCGAGTACCTTAAAGCGAAGTTTGAGCAAGCAGGATGGAGGTGTCCCGAGGAGTGA 1597
421 AATCGAGTACCTTAAAGCGAAGTTTGAGCAAGCAGGATGGAGGTGTCCCGAGGAGTGA 480
1598 ACAAAGCACTGAGCTTCTCCATCATATCAGAGAGAGGAGCGGAGAGCTCCAAAGAAATCA 1657
481 ACAAAGCTCTGAGCTTCTCCAGCATCCAGAGAGAGGAGCGGAGAGCTCCAGGAGATCA 540
1658 AAGAGCAGGAGTACCAGGCTCAAGTGGAGAAATAGGTTGATGATGAATCAGTTGGAAG 1717
541 AGGAGCAGGAGTACCAGGCTCAGTGGAGAGAGTGGAGCTGATGATGAATCAGTTGGAAG 600
1718 AGGATCTGTCTCAGCAAGAGAGCGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTC 1777
601 AAGACCTGGTGTGAGCCCGCAGAGCGAGCATCTCTACGAGTCTGAGCTGAGGAGTCTC 660
1778 GGCTTGTCTGAGAAATTCAGCGGAAACCGCAGAGAGTGTGAGCTGAGCTGAGGAGTCTC 1837
661 GGCTTGTCTGAGAAATTCAGCGGAAACCGCAGAGAGTGTGAGCTGAGCTGAGGAGTCTC 720
1838 CTAGGATCAAGGAGAGCTTGAAGTGGAGAAATATGCGAAATCTGGAGAAATCAATCTG 1897
721 CTAAGGACCAAGGAGAGCTTGAAGTGGAGAAATATTCGAATCTGGAGAAATCAATCTG 780
1898 AGCAGCAGCTCAAAATTCAGGAGCTCAAGAGAAACTGGAGAGAGTGTGAGAGGAGCGAG 1957
781 AGCAGCAGCTGAAGATCCAGAGCTCCAGGAGAGCTGGAGAGAGTGTGAGAGAGCGAGCA 840
1958 CCGAGAGGAGCTGGAGAGCTGAGAGCTGAGAGAGCTGAGAGAGGATTTCTTCTGAGGAGCATCAGAAAGA 2017
841 CAGAGGAGCTGAGAGCTTCTGAGAGAACTGAGAGAGCTTCTGAGAGAGCTTCTGAGAGAGCTGAGAG 2077
2018 AGCTGGTGGAGCTGAGAGAACTGAGAGAGCTTCTGAGAGAACTGAGAGAGCTTCTGAGAGAGCTGAGAG 2137
883 AGCAGAGAGCTGAGAGAGCTGAGAGAGCTTCTGAGAGAACTGAGAGAGCTTCTGAGAGAGCTGAGAG 927
2078 CCATGGAGCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTTCTGAGAGAGCTTCTGAGAGAGCTTCTGAGAG 2137
928 CCTCCGAGGAGCTCAAAAGAGAGCTGAGAGAGCTTCTGAGAGAGCTTCTGAGAGAGCTTCTGAGAGAGCT 962
2138 AGCAGAGTGGCTGATATAAATTTCTGAGAGCTCGAAGAGAAACATCGGAGAGCTTCTGAGAGCTTCTGAG 2197
963 -----GAGAGCTGGAGAGAGAAACATCGGAGAGCTTCTGAGAGCTTCTGAGAGCTTCTGAGAGCTTCTGAG 1002
2198 CCCAGCAGCTTGAAGTGCAGCTGAGAGAGAGAGAGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2257
1003 CTCACATCTAGAAGTACCTTGAAG 1062
2258 TGTGGAGAGTACAGATTAAG 2317
1063 TATTGGAGAGTACAGATTAAG 1122

QY	231.8	AGAGACACGAGGAGGAGGCCCATGAGAAGGCGCAAAATTCACAGCGAAACAGAAAGCGCATGA	2377
DB	112.3	AGAGACACGAGGAGGAGGCCCAACAGAGAGGCGAAGATCCTCAGCGACGACGAAAGCGCATGA	1182
QY	237.8	TCAATGCTATGATTTCCAAAGATCAGATCCCTCGGAACAGAGGATTTGTGGAACTGCTCTGAAG	2437
DB	119.3	TCACCCGATGATTTCCAAAGATCCGATCCCTCGGAGCAGAGATCGTGGAGCTGTCGGAAG	1242
QY	243.8	CCAATPAACTTTCGACCAATATGACGATCTTTTTACCCAAAGGAAACATGAAGGCCCAAGAAG	2497
DB	124.3	CCAACAAGCTTTCGCGCAAAACAGCAGTCTCTTACCCACGAGGAAACATGAAGGCCCGGAAG	1302
QY	249.8	AGATGATTTCTGAACTCAGGCAACAGAAATTTTACCTTGGAGACACAGGCTTGGGAAGTTGG	2557
DB	130.3	AGATGATCTCAGAACTCAGGCAGCAGAAATTTTACCTTGGAGACGACGGCCGGAAGCTGG	1362
QY	255.8	AGGCCCAGAAACCGAAAACTGGAGGACAGCTGGAGAAAGATCAGCCACCAAGACACACAGTG	2617
DB	136.3	AGGCCCAGAAACCGAAAGCTGGAGAGCAACTGGAGAGATCAGCCACCAAGATCAGATG	1422
QY	261.8	ACAGAAATCGGCTGCTGGAACCTGGAGACAAAGATTGCGGAGGTCAGTCTAGAGCAAGAGG	2677
DB	142.3	ACAAGAGTCGGCTCTGGAGCTGGAAACAAAGGCTGAGGGAGGTCAGCCTGGAGCAGGAGG	1482
QY	267.8	AGCAGAAACTGGAAGCTCAAGCGCCAGCTCACAGAGCTACAGAGCTCTCCCTCGACGAGCGCG	2737
DB	148.3	AGCAGAGCTGGAGCTAGAGCGGCAGCTCAGGAGCTGAGAGCTGCTCCCTCGAGAGCGCG	1542
QY	273.8	AGTCACAGTTGACAGCCTCGAGGCTGCACGGGCGGCTTGGAGAGCCAGCTTCGCGCAGG	2797
DB	154.3	AGTCCCAGCTGACGGCCCTCGAGGCCCGCGGCGAGCGCTGGAGAGCCAGCTCCGCCAGG	1602
QY	279.8	CGAAGACAGAGCTGGAAGAGACACACAGCAAGCTGAAGGAGAGATCCAGGCACTCAACGG	2857
DB	160.3	CGAAGACAGAGCTGGAGAGACAAACCGGAAGCGGAGGAGAGATCCAGGGCGCTCACCG	1662
QY	285.8	CACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAATCA	2917
DB	166.3	CACATCCGATGAGATCCAGCGCAAAATTCGATGCCCTTCGCAACAGCTGCACGGTCATCA	1722
QY	291.8	CAGACCTTGGAGGAGCAGCTAAACAGCTGACCGAGACAAAGCTGAACTCAACACCAAA	2977
DB	172.3	CCGACCTTGGAGGAGCAGCTGAAACAGCTCACCAGGACAAACGCCGAGCTCAACACCAAA	1782
QY	297.8	ACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCACACGACGAGATTGTACAAC	3037
DB	178.3	ACTTCTACTGTCCAAACAACTCGATGAGGCTTCCGGGGCCCAATGACGAGATTGTGCAGC	1842
QY	303.8	TGCGAAGTGAAGTGGACCATCTCCGCGGGAGATCAAGGAAACGAGAGATGCAAGCTTACCA	3097
DB	184.3	TGCGGAGTGAAGTGGACCATCTCCGCGGTGAGATCAAGGACGGGAGATGCAAGCTCACCA	1902
QY	309.8	GCCAGAAGCAAAACGATGGAGGCTCTGAAGACCAAGCTGCACCATGCTGGAGGAAACAGGTCA	3157
DB	190.3	GCCAGAAGCAAAACGATGGAGGCTCTGAAGACCAATGCAACATGTCACATGTTGGAGAGCAGGTC	1962
QY	315.8	TGGATTTGGAGGCCCTAAACGATGAGCTGTGAAGAAAAGCGGCGAGTGGGAGGCTCGA	3217
DB	196.3	TGGACCTTGGAGGCCCTGAACGACGAGCTGCTGGAGAAGGAGGCCAGTGGGAGGCTGCG	2022
QY	321.8	GGAGGCTCTCTGGGTGATGAGAAATCCAGTTCGATGTCGGGTTCGAGAGCTGCAGAGNA	3277
DB	202.3	GGAGGCTCTCTGGCGACGAGAGTCCAGTTGAGTTCGAGTTCGAGTTCGAGCTTACAGAGNA	2082
QY	327.8	TGCTGACACCGAGAAACACAGACAGGCGCAGACCGGATCAGCGGATCACCGAGTCTCGCC	3337
DB	208.3	TGCTGACACCGAGAAACACAGACAGGCTGAGATTCCTCCGCTCGACGAGGCTC	3397
QY	333.8	AGGTGCTGAGGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCCTCCGCTCGACGAGGCTC	3457
DB	214.3	AGGTGCTGAGGTTGGCGGTGAAGGAACACAGGCCGAGATTCCTGCTCTGCAGCAGGCTC	2202
QY	339.8	TCAAAGACAGAGCTGAAGGCCGAGAGGCTCTCTCAAGCTCAATGACCTGGAGAGAAGA	3457

Db	2203	TCAGGAGCAGAAGCTCAAGCCGAGAGCCTGTGCGACAAGCTCAACGACCTGGAGAAGA	2262
Qy	3458	AGCATGCTATCTGTTGAATGAATGCCCGAGCTTTACAGCAGAGAGCTGGAGACTGAACGAG	3517
Db	2263	AGCAGCCCATCTCGAGATGAACCCCGGAGCTTTACAGCAGAACTTAGACACAGACGGGG	2322
Qy	3518	AGCTCAAAACAGAGGCTTCTGGAAGACAAGCAAAATTCACAGCAGCAGATGGAACCTGCAGA	3577
Db	2323	AGCTCAAAACAGAGGCTTCTGGAGGAGCAAGCAGAAATTGACGACGAGATGGAACCTGCAGA	2382
Qy	3578	AAAATCACATTTTCGCTGATCTCAAGGACTGCAAGAGCTCTAGATTCGGGCTGATCTAC	3637
Db	2383	AGAAACCAATCTTCAGACTACGCAAGGGCTGAGGAGGCGCTGGACCGGCGGATCTGC	2442
Qy	3638	TGAAGACAGAAAAGAGTACTTGGAGTATCAGCTGGAAACATTTACAGTTCTCTATTTCTC	3697
Db	2443	TGAAGACAGAAAAGAGGCGACTGGAGTACAGCTGGAAGAACATTCAGTTCTCTACTCTC	2502
Qy	3698	ATGAAAAGGTGAAAATGGAAGGCATATTCTCAACAAACAAACCTCATTTGATTTCTGC	3757
Db	2503	ACGAGAAGTGAAAATGGAAGGCACATCTCTCAGCAAAACAAACTCATTTGATTTCTGC	2562
Qy	3758	AAGCCAAAATGGACCAACTGCTTAAAGAGAAAAGGGTTTATTAGTCGACGGAAAGAGG	3817
Db	2563	ANGCCAAAATGGACCAAGCTGCTTAAAGAGAAA-----	2595
Qy	3818	ACCCTGCTTTACCCACACAGGTTCTCTGCGATACAATGAGCTGAAGCTGGCCCTGGAGA	3877
Db	2596	-----AGGTTCTCTGCGTACAAATGAGCTGAAGCTAGCCCTGGAAA	2637
Qy	3878	AGGAGAAGCTCGCTGTGACAGAGTACAGGAAGCCCTTCAGAAAGACCCGATCGAGCTCC	3937
Db	2638	AGGAGAAGCCCGATGCGCGGAGCTGGAGAGGCCCTTCAGAAAGACCCGATCGAACTCC	2697
Qy	3938	GGTCGCGCCGGGAGGAGCTGCCCAACGGAAACGACGACCAACCCACACCCCATCCAGC	3997
Db	2698	GCTCTGCGCGGAGGAGCTGCCCAACGGAAACGACGACCAACCCGATCGCTACGC	2757
Qy	3998	CAGCCACCGGAGGACAGATCGCATGTCCGCATCTGTGGGTGCCAGAGCAACAGC	4057
Db	2758	CAGCCACTGGGAGGACAGATCGCATGTCTGCCATTTGTGGGTGCCCGAGCAACAGC	2817
Qy	4058	CCAGTGCCATGAGCTGCTGGCCCGCCCATCCAGCGCAGAAAGAGTCTTCAACTCCAG	4117
Db	2818	CCAGTGCCATGAGCTGCTGGCCCAACATCCAGCGCAGAAAGAGTCTTCAACTCCAG	2877
Qy	4118	AGAAATTAAGTCGGGCTTTAAGAAACGATGACACACAATATTCTCAACGATTCACG	4177
Db	2878	AGAAATTCAGCCGGCTCTGAAGAGCGCATGCCACAACTCCCTACCCGGTTTAATG	2937
Qy	4178	TAGACTTGAACATGCGAGCACAAAGTGTGCTGTGCTGATACCGTGCACTTTTGAC	4237
Db	2938	TGGGCTTGAACATGAGAGCAACAGTGCCTGTGCTGTGATTAATGTGCACTTTTGAC	2997
Qy	4238	GCCAGGCATCCAAATGCTTCGAATGTCAGGTGATGTGCACCCCAAGTGTCTCACGTGT	4297
Db	2998	GCCAGGCATCCAAATGCCTTCGAATGTCAAGTGTGTGCACCCCAAAATGCTCCACCTGT	3057
Qy	4298	TGCCAGCCCTGCGGCTGCTGCTGAATATGCCACACACTTCCAGGAGGCTTTCTGC	4357
Db	3058	TGCCCGCTACCTGTGGCTTGCACAGTGAATAATGCCACACACTTCACTGAGGCTTTCTGC	3117
Qy	4358	GTGACAAAATGAATCCCCAGGTCTCCAGCAACAGGAGCCAGCAGCAGCTTGCACTGG	4417
Db	3118	GGGACAAAATGAATCCCGGGCTCCAGAGCAAGGAGCTTGGCAGCAGCTTGCACTGG	3177
Qy	4418	AGGGTGGATGAAGTGGCCAGGATTAACAAACGAGGACAGCAAGCTTGGACAGGAAGT	4477
Db	3178	AGGGTGGATGAAGTGGCCAGGATTAACAAACGAGGAGCAGCAAGCTTGGACAGGAAGT	3237
Qy	4478	ACATTTCTCGAGGAGTCAAAAGTCTCAATTTATGACAAATGAAGCCAGAGAACTGGAC	4537

```
Db 3238 ACATTGTTCTGGAGGGTCAAAAGTCTCTCATCTATGACATGAAGCCAGAGAAGCTGGAC 3297
Qy 4538 AGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCGACGGGGATGTATCTATTCATGGTG 4597
Db 3298 AGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCGACGGGGATGTATCTATTCATGGTG 3357
Qy 4598 CCGTTGTTGCTTCCGAATCTCGAATACAGCAAAAGCA----- 4635
Db 3358 CCGTTGTTGCTTCCGAATCTCGAATACAGCAAAAGCAAGATGTCCCATACATCTCGAAGA 3417
Qy 4636 ----- 4635
Db 3418 TGGAGTCTCACCCACACACACCTGCTGGCCCTGGGAGAACCTCTACTTGTGTAGCACCCA 3477
Qy 4636 ----- 4635
Db 3478 GCTTCCCCGCAACAGCAGCGCTGGGTTCACCGCCTTAGAATCTGTGTCGTCAGGTGGGAGAG 3537
Qy 4636 -----GAAAAAGCAGAAGTGTATGCTAAACTGCTTGGAAACTCCCTGTGTGAATCTGG 4687
Db 3538 TTTCTAGGGAAGAGCCGAAGCCGATGCTAAATTAATTTGGAATCTCTCTGCTGAAACTGG 3597
Qy 4688 AAGGTGATGACCGTCTAGACATGAATGCACTGCACTGCTTCAAGTACAGCAGGTGTGTGG 4747
Db 3598 AGGGCGATGACCGGCTTGACATGAATGCACTGCACTGCTTCAAGTACAGCAGGTGTGTGG 3657
Qy 4748 TGGGCACCGAGGAAGGGTCTTACGCGCTGAATGTCTTGAATAACTCCCTTAACCCATGTCC 4807
Db 3658 TGGGCACCGAGGAAGGGTCTTACGCGCTGAATGTCTTGAATAACTCCCTTAACCCATGTCC 3717
Qy 4808 CAGGAATTTGAGCAGTCTTCCAAATTTATATATATCAAGGAACCTGGAGAACTACTCATGA 4867
Db 3718 CAGGAATTTGAGCAGTCTTCCAAATTTATATATATCAAGGAACCTGGAGAACTGTCTCATGA 3777
Qy 4868 TAGCAGGAAGAGCGGGACCTGTCTTGTGAGCTGAAGAAAGTGAACAGTCCCTGG 4927
Db 3778 TAGCAGGAAGAGCGGGCTCTGTGCTGTGAGCTGAAGAAAGTGAAGAGTGAAGAGTCCCTGG 3837
Qy 4928 CCCAGTCCACCTGCTGCTGCCAGCCGACATCTCAACCAACATTTTGAAGCTGCAAG 4987
Db 3838 CGCAGTCAACCTGCTGCTGCCAGCCGACGTCTCCCCAACATATTCGAAGCCGTCAAAG 3897
Qy 4988 GCTGCCACTTTGTTGGGGGAGCAAGATTTGAGAACCGGCTCTGCATCTGTGACGCCATGC 5047
Db 3898 GCTGCCACTTTGTTGCTGTGTCAGAGATCGAAGACAGCCTGTGCAATCTGCGCCGTATGC 3957
Qy 5048 CCAGCAAGCTGCTCATTTCTCCGCTACACAGAAACCTCAGCAAAATACTGCAATCCGGAAG 5107
Db 3958 CAAGCAAGTGTCTATCTCCGCTACAAATGACAACTCAGCAAGTACTGCAATCCGCAAG 4017
Qy 5108 AGATAGAGACCTCAGAGCCCTGAGCTGTATCCATTCACCAATTAAGTATCTCATTTG 5167
Db 4018 AGATCAGAGACCTCAGAGCCCTGAGCTGTATCCATTTCAACCACTACAGCATCTCATTTG 4077
Qy 5168 GAACCAATTAATCTACGAATTCGACATGAAGCAGTACAGCTCGAGGAATTCCTGGATA 5227
Db 4078 GGAACCAATTAATCTATGAGATCGACATGAAGCAGTACAGCTTGTGATGATTTCTTGGACA 4137
Qy 5228 AGAATGACCAATTCCTTGGCACCTGTGTGTTGTCGCGCTCTTCCACAGCTTCCCTGTCT 5287
Db 4138 AGAACGACCAATTCCTTGGCACCTGTGTGTTGTCGCTCTCTCGTCCAACAGCTTCCCTGTCT 4197
Qy 5288 CAATGCTGAGGTGAACAGCGGAGGCGAGGAGAGTACTGTCTGTGTTTCCAGCAAT 5347
Db 4198 CCATTTGACGGGCAACAGCGCGGGGAGCGAGGAAGAAATACCTGCTGTGCTTCCAGCAAT 4257
Qy 5348 TTGAGGTGTTCTGTTGATTTCTTACGGAAGAGCTAGCCGACAGATGATCTCAAGTGGAGTC 5407
Db 4258 TTGGGTGTTCTGTTGATTTCTTACGGAAGAGTGAAGCGGAGAGTATCTCAAGTGGAGTC 4317
Qy 5408 GCTTACCTTTGGCTTTGGCTTACAGAGAACCTTATCTGTTTGTGACCCACTTCAACTCAC 5467
Db 4318 GCTTACCTTTGGCTTTGGCTTACAGAGAACCTTATCTGTTTGTGACTCACTTCAACTCCC 4377
```

```
Qy 5468 TCGAAGTAATTGAGATCCAGGCACGCTCTCTCAGCAGGAGACCCCTGCCGAGCGTACCTGG 5527
Db 4378 TCGAAGTCAATTGAGATCCAGGCACAGATCTCACTGGGGAGCCCTGCCGAGCATATCTGG 4437
Qy 5528 ACATCCCGAAACCCGCGCTACCTGGGGCCTTGCATTTCTCTCAGGAGCGATTTACTTGGCGT 5587
Db 4438 AAATTCAAAACCTCTCTACCTGGGGCCTTGCATTTCTCTCCGAGCGATTTACTTGGCGT 4497
Qy 5588 CCTCATACAGGATAAATTAAGGGTCAATTTGTGCAAGGGAACCTCTGTAAGGAGTCCG 5647
Db 4498 CCTCATACAGGACAAGTTAAAGGTCATATGTCGCAAGGAAACCTCTGTAAGGAGTCA 4557
Qy 5648 GCACTGAACACCAACCGGGGCGCTTCCACTTCCGCGAGCGCCCCAACAGCGAGGCCCCAC 5707
Db 4558 GCACTGAGCAGCACCGGGTGTCTTCCACTTCCGCGAGCGCCCCAACAGCGAGGCCCCAC 4617
Qy 5708 CCAGCTACACAGCAGCATCAACCAAGCGGTGCTTCCAGCCAGCGCCCGCCCGAAGGCC 5767
Db 4618 CAACATACAAAGAGCAGCATCAACCAAGCGCACACCGCTACCGAGACAGAGAGGTCGAGAC 4677
Qy 5768 CCAGCCACCCCGCAGAGCCAAAGCACACCGCTACCGAGACAGAGAGGTCGAGAC 5821
Db 4678 CCAGCCACCCCGCAGAGCCAAAGCACACCGCTACCGAGACAGAGAGGTCGAGAC 4737
Qy 5822 AGCTGCGCAGGACAAGTCTCTTGGCGGCCCTTGGAGCGAGAGAGTCCCCCGGCGGA 5881
Db 4738 AGCTGCGCAGGACAAGTCTCTTGGCGGCCCTTGGAGCGGAGAGTCCCCAGGCCGA 4797
Qy 5882 TGCTCAGCACCGCAGAGAGAGCGGTCCCCCGGAGGCTGTTTGAAGACAGCAGCAGGGGCC 5941
Db 4798 TGCTCAGCATAGGAGAGAGCGGTCCCCAGGGAGACTGTTTGAAGACAGCAGCAGGGGCC 4857
Qy 5942 GGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCTCCAGGTGAACAAGGGAAGAGGGGAGA 6001
Db 4858 GGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCTCCAGGTGAACAAGGGAAGAGGGGAGA 6017
Qy 6002 GTGCTCTCAAGTTT 6016
Db 4918 CTTCAGTATAGTCT 4932
```

Search completed: July 3, 2004, 16:52:14
Job time : 15441 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 12:14:14 ; Search time 9465 Seconds
(without alignments)
19431.714 Million cell updates/sec

Title: US-10-017-216-3

Perfect score: 6159

Sequence: 1 atgttgagttcaaatagg.....ttctgagacagattatgc 6159

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_estcl.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rtd.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gasl.*

29: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	915.2	14.9	1011	13 B0070955	B0070955 AGENCOURT
2	815.6	13.2	966	13 B0071141	B0071141 AGENCOURT
3	810.2	13.2	830	13 B0181633	B0181633 AGENCOURT
4	750.6	12.2	849	12 B1253509	B1253509 602937370

5	716.4	11.6	920	13	BX342268	BX342268
6	692	11.2	879	13	BQ228524	BQ228524 AGENCOURT
7	688.8	11.2	958	11	BC031156	Mus muscu
8	629.2	10.2	652	10	BF905370	IL3-MT026
9	606.5	9.8	881	12	BG976452	602846269
10	589.8	9.6	746	14	CA749290	UI-M-FY0-
11	584.6	9.5	843	12	BG912161	BG912161 602812833
12	573.6	9.3	647	14	CF744580	UI-M-GV0-
13	572.8	9.3	730	14	CF723360	UI-M-GV0-
14	566.4	9.2	1085	12	BM904785	AGENCOURT
15	556.8	9.0	640	13	BQ807302	NISC_KK01
16	511.8	8.3	775	13	BUI05833	603005490
17	511	8.3	804	13	BUI236817	603411670
18	510.4	8.3	933	13	BQ232508	603408272
19	508.6	8.3	547	14	CF744826	UI-M-GV0-
20	491.6	8.0	875	13	BX723780	EX723780
21	490.8	8.0	875	13	BU900029	AGENCOURT
22	484.2	7.9	820	13	BQ941506	AGENCOURT
23	473	7.7	647	10	BQ207065	BB207065
24	472.6	7.7	657	13	BQ831488	LI61n2149
25	471.6	7.7	698	14	CD348416	UI-M-FY0-
26	466.6	7.6	681	10	AW605350	QV3-DT004
27	451.4	7.3	459	13	BU429033	UI-HF-BN0
28	450.6	7.3	484	10	BE834307	RC1-OT008
29	445.5	7.2	891	12	BI558919	603240969
30	443.5	7.2	600	12	BG808196	2082-86 M
31	443.2	7.2	719	13	BUS39418	603515216
32	441.4	7.2	513	14	CF361392	827723 MA
33	437.8	7.1	860	14	CF223944	AGENCOURT
34	437.6	7.1	589	10	BF819093	QV2-C1010
35	435.2	7.1	1075	13	BQ425970	AGENCOURT
36	434.5	7.1	596	14	CF540358	UI-M-GV0-
37	431.2	7.0	497	10	BE909486	601502838
38	424.2	6.9	662	13	BU104008	603004106
39	417.4	6.8	427	12	BG944539	axs51111.x
40	417.2	6.8	749	13	BU364866	603567832
41	415	6.7	678	13	BU403714	604137885
42	405.6	6.6	617	13	BQ417810	1k51b06.Y
43	398	6.5	936	12	BI758420	603026948
44	387.8	6.3	730	14	CA380902	660300 NC
45	384.2	6.2	1013	13	BUI08993	603110947

ALIGNMENTS

RESULT 1
BQ070955
LOCUS BQ070955 1011 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6855647 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5923441
5', mRNA sequence.

ACCESSION BQ070955 GI:19900001

VERSION BQ070955

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1011)

1 (bases 1 to 1011)

NH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnl.gov>

Plate: L1CM2091 row: p column: 02

High quality sequence stop: 634.

FEATURES		Location/Qualifiers	
QY	source	1. .1011	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5923441"	
QY		/tissue_type="neuroblastoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC 47"	
		/notes="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."	
ORIGIN			
Query Match		14.9%; Score 915.2; DB 13; Length 1011;	
Best Local Similarity		97.9%; Pred. No. 4.6e-206;	
Matches 946; Conservative 0; Mismatches 17; Indels 3; Gaps 2;			
QY	2653	CGGAGGTCAGCTAGACACGAGGAGCAGAACTGGAGCTCAAGCGCGAGCTCACAGAG	2712
Db	1	CGGAGGTCAGCTAGACACGAGGAGCAGAACTGGAGCTCAAGCGCGAGCTCACAGAG	60
QY	2713	CTACAGCTCTCCCTGAGGAGCGCGAGTCACAGTTCACAGCCCTGCAGGCTGCAGGGCG	2772
Db	61	CTACAGCTCTCCCTGAGGAGCGCGAGTCACAGTTCACAGCCCTGCAGGCTGCAGGGCG	120
QY	2773	GCCTTGAGAGCGAGCTTCGCAGCGGGAAGACAGAGCTGGAAGACACACAGAGAAGCT	2832
Db	121	GCCTTGAGAGCGAGCTTCGCAGCGGGAAGACAGAGCTGGAAGACACACAGAGAAGCT	180
QY	2833	GAAGAGGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAAAATTCATGCT	2892
Db	181	GAAGAGGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAAAATTCATGCT	240
QY	2893	CTTCGTAACAGCTGTACTTAATACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG	2952
Db	241	CTTCGTAACAGCTGTACTTAATACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG	300
QY	2953	GACAGCGTGAACCTCAACACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT	3012
Db	301	GACAGCGTGAACCTCAACACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT	360
QY	3013	GGCGCCAACGACGAGATTGTCAACTCGCAAGTGAAGTGAGCCATCTCCGCGGAGATC	3072
Db	361	GGCGCCAACGACGAGATTGTCAACTCGCAAGTGAAGTGAGCCATCTCCGCGGAGATC	420
QY	3073	ACGGAACGAGATGACAGCTTACCGCCAGAGACAAACGATGGAGGCTCTGAAGACCG	3132
Db	421	ACGGAACGAGATGACAGCTTACCGCCAGAGACAAACGATGGAGGCTCTGAAGACCG	480
QY	3133	TGCACCATCTCTGAGGAACAGCTCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA	3192
Db	481	TGCACCATCTCTGAGGAACAGCTCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA	540
QY	3193	AAAGAGCGCAGTGGAGCGCTTGGAGGAGCGTCTCTGGGTGATGAGAAATCCAGTTTGA	3252
Db	541	AAAGAGCGCAGTGGAGCGCTTGGAGGAGCGTCTCTGGGTGATGAGAAATCCAGTTTGA	600
QY	3253	TGTCGGGTTTCGAGAGCTGCAGAGATGCTGGACCGAGAACAGACAGAGCGCGAGGCC	3312
Db	601	TGTCGGGTTTCGAGAGCTGCAGAGATGCTGGACCGAGAACAGACAGAGCGCGAGGCC	660
QY	3313	GATCAGCGGATCACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGAGCACAGGCT	3372
Db	661	GATCAGCGGATCACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGAGCACAGGCT	720
QY	3373	GAGATTCTCGTCTGCAGCAGGCTCTCAAAGAGCAGAGCTGAAGCGCGAGAGCTCTCT	3432

Db	721	GAGATTCTCGTCTGCAGCAGGCTCTCAAAGAGCAGAACTGGAGCGCGAGAGCTCTCT	780
QY	3433	GACAAGCTCAATGACCTGGAGAGACATGCTATGCTTGAATGAATGCC-CAAGCTT	3491
Db	781	GACAAGCTCAATGACCTGGAGAGACATGCTATGCTTGAATGAATGCCCGGAAGCTT	840
QY	3492	ACAGCAGAAGCTGGAGACTGAACGAGAGCTCAACACGAGAGCTTCTTGAAGACAGCCAA	3551
Db	841	ACAGCAGAAGCTGGAGACTGAACGAGAGCTCAACACGAGAGCTTCTTGAAGACAGCCAA	900
QY	3552	ATACACACAGATGACCTGCAGAAAAATCACATTTT--CGTCTGATCAAGGACTG	3609
Db	901	ATACACACAGATGACCTGCAGAAAAATCACATTTTTCGCGCTGGACTCAAGGACTG	960
QY	3610	CAAGAA 3615	
Db	961	CCAGAA 966	
RESULT 2			
BQ071141			
LOCUS			
DEFINITION		BQ071141 956 bp mRNA linear EST 02-APR-2002	
		AGENCOURT 6853098 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927502	
		5', mRNA sequence.	
ACCESSION		BQ071141	
VERSION		BQ071141.1 GI:19900187	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 956)	
AUTHORS		NIH-MGC http://mgc.nci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: ggapbs-rc@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LCM2102 row: i column: 07 High quality sequence stop: 650. Location/Qualifiers 1. .956 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5927502" /tissue_type="neuroblastoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC 47" /notes="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."	
FEATURES		source	
Query Match		13.2%; Score 815.6; DB 13; Length 956;	
Best Local Similarity		97.1%; Pred. No. 2.2e-182;	
Matches 874; Conservative 0; Mismatches 19; Indels 7; Gaps 4;			
QY	2653	CGGAGGTCAGTCTAGAGCAGCAGGAGCAGAACTGGAGCTCAAGCGCGAGCTCACAGAG	2712


```
Db 1 CGGGAGGTCAGTCTAGACGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAG 60
QY 2713 CTACAGCTCTCCCTCAGGAGCGCGAGTCACAGTTGACAGCCCTGACGGCTGCACGGCG 2772
Db 61 CTACAGCTCTCCCTCAGGAGCGCGAGTCACAGTTGACAGCCCTGACGGCTGCACGGCG 120
QY 2773 GCCCTGGAGAGCCAGCTTCGCCAGGCGGAGACAGAGCTGGAGAGACACAGCAGAGCT 2832
Db 121 GCCCTGGAGAGCCAGCTTCGCCAGGCGGAGAGCAGAGCTGGAGAGACACAGCAGAGCT 180
QY 2833 GAAGAGGAGATCCAGGCACTCACGCGCACATAGATGAAATCCAGCGCAAAATTTGATGCT 2892
Db 181 GAAGAGGAGATCCAGGCACTCACGCGCACATAGATGAAATCCAGCGCAAAATTTGATGCT 240
QY 2893 CTTCTACAGCTGTAATGTAATCAAGACCTGGAGAGAGCTTAAACAGCTGACCGAG 2952
Db 241 CTTCTACAGCTGTAATGTAATCAAGACCTGGAGAGAGCTTAAACAGCTGACCGAG 300
QY 2953 GACAAAGCTGAACTCAACCAACCAAACTTCTACTTGTCCAAACCACTCGATGAGGCTTCT 3012
Db 301 GACAAAGCTGAACTCAACCAACCAAACTTCTACTTGTCCAAACCACTCGATGAGGCTTCT 360
QY 3013 GCGGCAACGACGAGATGTAACACTGCGAAGTGAAGTGGACCAATCTCCGCGGAGATC 3072
Db 361 GCGGCAACGACGAGATGTAACACTGCGAAGTGAAGTGGACCAATCTCCGCGGAGATC 420
QY 3073 ACGGACGAGATCCAGCTTACCGAGCAGAGCAAAAGTGGAGGCTCTGAGACCAAG 3132
Db 421 ACGGACGAGATCCAGCTTACCGAGCAGAGCAAAAGTGGAGGCTCTGAGACCAAG 480
QY 3133 TGCACCATGCTGGAGGAACAGCTCATGATTTGGAGGCGCCCTAAACGATGATGCTGAGAA 3192
Db 481 TGCACCATGCTGGAGGAACAGCTCATGATTTGGAGGCGCCCTAAACGATGATGCTGAGAA 540
QY 3193 AAAGAGCGCAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTGAG 3252
Db 541 AAAGAGCGCAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTGAG 600
QY 3253 TGTCTGGTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3312
Db 601 TGTCTGGTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 660
QY 3313 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGAGTGAAGAGCAGCAGAGCT 3372
Db 661 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGAGTGAAGAGCAGCAGAGCT 720
QY 3373 GAGATCTCGCTCTCAGCAGGCTCTCAAGAGCAGAGCTGAGGCGGAGGCTCTCT 3432
Db 721 GAGATCTCGCTCTCAGCAGGCTCTCAAGAGCAGAGCTGAGGCGGAGGCTCTCT 780
QY 3433 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAATGAAATGCGCGAAGC 3489
Db 781 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAATGAAATGCGCGAAGC 840
QY 3490 TTACAGAGAA--GCTGGAGACTGAAC--GAGAGCTCAACAGAGCTTCTGGAGAGCA 3545
Db 841 TTACAGAGAAAGCTGGGAGACTGAACCGAGAGCTCAACAGAGGCTTCTGGAGAGCA 900
```

RESULT 3

```
BU181633
LOCUS BU181633
DEFINITION BU181633 830 bp mRNA linear BST 04-SEP-2002
5', mRNA sequence.
ACCESSION BU181633
VERSION BU181633.1 GI:22695617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
```

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM13459 row: e column: 19
High quality sequence stop: 652.
Location/Qualifiers
1. 830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6140538"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN

```
Query Match 13.2%; Score 810.2; DB 13; Length 830;  
Best Local Similarity 99.4%; Pred. No. 3.9e-181;  
Matches 823; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 2387 TGGATTCCTCAAGATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAGCCCAATAAAC 2446  
Db 1 TGGATTCCTCAAGATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAGCCCAATAAAC 60  
QY 2447 TTGAGCAAAATAGCAGCTTTTATCCCAAAGAAATGAAAGGCGCCCAAGAGAGATGATTT 2506  
Db 61 TTGAGCAAAATAGCAGCTTTTATCCCAAAGAAATGAAAGGCGCCCAAGAGAGATGATTT 120  
QY 2507 CTGAACTCAGCAGCAAGAAATTTTACCTGGAGACAGAGCTGGAGCTGGAGGCGCCAG 2566  
Db 121 CTGAACTCAGCAGCAAGAAATTTTACCTGGAGACAGAGCTGGAGCTGGAGGCGCCAG 180  
QY 2567 ACCGAAATCTGGAGGAGCAGCTGGAGAGATCAGCCCAAGACCAAGACCAAGATC 2626  
Db 181 ACCGAAATCTGGAGGAGCAGCTGGAGAGATCAGCCCAAGACCAAGATC 240  
QY 2627 GGCTGCTGGAACTGGAGCAAGATTGGGGAGGTGAGTCTAGAGCAGGAGCAGAGAAC 2686  
Db 241 GGCTGCTGGAACTGGAGCAAGATTGGGGAGGTGAGTCTAGAGCAGGAGCAGAGAAC 300  
QY 2687 TGGAGCTCAAGCGCCAGCTCAGAGACTACAGCTCTCCCTGCGAGGAGCGGAGTCAAGT 2746  
Db 301 TGGAGCTCAAGCGCCAGCTCAGAGACTACAGCTCTCCCTGCGAGGAGCGGAGTCAAGT 360  
QY 2747 TGACAGCCCTGCGAGGTGCGACGGGCGGCCCTGGAGAGCAGCTTTGGCCAGCGGAGAGCAG 2806  
Db 361 TGACAGCCCTGCGAGGTGCGACGGGCGGCCCTGGAGAGCAGCTTTGGCCAGCGGAGAGCAG 420  
QY 2807 AGCTGGAGAGACCAAGCAGAGAGAGCTCAAGAGAGAGTCCAGGCACTCAGCGCACATAGAG 2866  
Db 421 AGCTGGAGAGACCAAGCAGAGAGAGCTCAAGAGAGAGTCCAGGCACTCAGCGCACATAGAG 480  
QY 2867 ATGAAATCCAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCACAGAGCTGG 2926  
Db 481 ATGAAATCCAGCGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCACAGAGCTGG 540  
QY 2927 AGGAGAGCTTAAACAGCTGACCGAGAGCAACCGTGAACCTCAACCAACCAAACTTCTACT 2986  
Db 541 AGGAGAGCTTAAACAGCTGACCGAGAGCAACCGTGAACCTCAACCAACCAAACTTCTACT 600  
QY 2987 TGTCCAAACCAACTCGATGAGGCTTCTGGCGCCCAACGAGAGAGTTGTACAACTGCGAAGTG 3046
```

TITLE
JOURNAL
COMMENT

```

Db      601  TGTCCAAACACTCGATGAGGCTTCTGGGCCCAACGACGAGATTGTACAACTGCGAGTG 660
Qy      3047  AATGAGCAATCTCCGCGGAGATCACGGAACGAGAGATGAGCTTACCAAGCAGAGC 3106
Db      661  AATGAGCAATCTCCGCGGAGATCACGGAACGAGAGATGAGCTTACCAAGCAGAGC 720
Qy      3107  AATGAGCAATCTCCGCGGAGATCACGGAACGAGAGATGAGCTTACCAAGCAGAGC 3165
Db      721  AATGAGCAATCTCCGCGGAGATCACGGAACGAGAGATGAGCTTACCAAGCAGAGC 780
Qy      3166  GAGGCGCTTAACGATGAGCTGTAGAAAAGAGCGGCGAGTGGAGGCC 3213
Db      781  GAGGCGCTTAACGATGAGCTGTAGAAAAGAGCGGCGAGTGGAGGCC 828

RESULT 4
BI253509 849 bp mRNA linear EST 17-JUL-2001
LOCUS 602973370F1 NIH_MGC_12 Homo sapiens cdna clone IMAGE:5112737 5',
DEFINITION mRNA sequence.
ACCESSION BI253509
VERSION BI253509.1 GI:14805003
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 849)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11273 row: 1 column: 18
High quality sequence start: 2
High quality sequence stop: 772.
Location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5112737"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

FEATURES
source
Query Match 12.2%; Score 750.6; DB 12; Length 849;
Best Local Similarity 98.5%; Pred. No. 6.1e-167;
Matches 789; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Qy      4659  ACTGCTTGAAACTCCCTGAACTGCAAGTGATGACCGCTCTAGACATCAACTGCAC 4718
Db      3 ACAGCGCGAAACTCCCTGCTGAACTGGAAGTGATGACCGCTCTAGACATCAACTGCAC 62
Qy      4719  GCTGCCCTTCAGTGACAGGTGCTGTGTGGGCGACCGAGGAAGGCTCTAGCGCCCTGAA 4778
Db      63 GCTGCCCTTCAGTGACAGGTGCTGTGTGGGCGACCGAGGAAGGCTCTAGCGCCCTGAA 122
Qy      4779  TGTCTTGAAAACTCCCTAACCATGTCTCCAGGAATTGAGCAGTCTTCCAAATTATAT 4838

```

```

Db      123  TGTCTTGAAAACTCCCTAACCATGTCTCCAGGAATCGGAGCAGTCTTCCAAATTATAT 182
Qy      4839  TATCAAGACCTGAGAAAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTGTCTTGT 4898
Db      183  TATCAAGACCTGAGAAAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTGTCTTGT 242
Qy      4899  GGAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCCGACAT 4958
Db      243  GGAAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCCGACAT 302
Qy      4959  CTCAACCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGGCAAGATTGA 5018
Db      303  CTCAACCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGGCAAGATTGA 362
Qy      5019  GAACGGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTGTCTCATTTCTCGGCTACAACGA 5078
Db      363  GAACGGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTGTCTCATTTCTCGGCTACAACGA 422
Qy      5079  AAACCTCAGCAAACTACTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTAT 5138
Db      423  AAACCTCAGCAAACTACTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTAT 482
Qy      5139  CCAGTTCAACCAATACAGTATCTCTTGGAAACCAATTAATTTACGAATTCACATGAA 5198
Db      483  CCAGTTCAACCAATACAGTATCTCTTGGAAACCAATTAATTTACGAATTCACATGAA 542
Qy      5199  GCAGTACACGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGGCACCTCTGTGTT 5258
Db      543  GCAGTACACGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGGCACCTCTGTGTT 602
Qy      5259  TGCGGCTCTTCCAAACAGCTTCCCTGTCTCAATGTGCGAGTGAAACAGCGCAGGCGAGCG 5318
Db      603  TGCGGCTCTTCCAAACAGCTTCCCTGTCTCAATGTGCGAGTGAAACAGCGCAGGCGAGCG 662
Qy      5319  AGAGGAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTCTCGTGGATTTTACGGAAGACG 5378
Db      663  AGAGGAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTCTCGTGGATTTTACGGAAGACG 722
Qy      5379  T-AGCCGACAGAGATCTCAAGTGAGTGCCTTACCTTTGGCCTTTGCTTACAGAGAAC 5437
Db      723  TAAAGCCGACAGAGATCTCAAGTGAGTGCCTTACCTTTGG-CTTTGCCCTACCGAGAA- 780
Qy      5438  CCTATCTGTTTGTGACCCACT 5458
Db      781  CCTATCTGTTTGGACCCACT 801

RESULT 5
BX342268 920 bp mRNA linear EST 02-MAY-2003
LOCUS BX342268
DEFINITION CDNA clone CSDBK011YB20 5-PRIME, mRNA sequence.
ACCESSION BX342268
VERSION BX342268.1 GI:30334095
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 920)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6533.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSDBK011DAL00P1&cluster=6533.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

```

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK011DA10QPL.

FEATURES	SOURCE
1. Feature 1	Source 1
2. Feature 2	Source 2
3. Feature 3	Source 3
4. Feature 4	Source 4
5. Feature 5	Source 5
6. Feature 6	Source 6
7. Feature 7	Source 7
8. Feature 8	Source 8
9. Feature 9	Source 9
10. Feature 10	Source 10
11. Feature 11	Source 11
12. Feature 12	Source 12
13. Feature 13	Source 13
14. Feature 14	Source 14
15. Feature 15	Source 15
16. Feature 16	Source 16
17. Feature 17	Source 17
18. Feature 18	Source 18
19. Feature 19	Source 19
20. Feature 20	Source 20
21. Feature 21	Source 21
22. Feature 22	Source 22
23. Feature 23	Source 23
24. Feature 24	Source 24
25. Feature 25	Source 25
26. Feature 26	Source 26
27. Feature 27	Source 27
28. Feature 28	Source 28
29. Feature 29	Source 29
30. Feature 30	Source 30
31. Feature 31	Source 31
32. Feature 32	Source 32
33. Feature 33	Source 33
34. Feature 34	Source 34
35. Feature 35	Source 35
36. Feature 36	Source 36
37. Feature 37	Source 37
38. Feature 38	Source 38
39. Feature 39	Source 39
40. Feature 40	Source 40
41. Feature 41	Source 41
42. Feature 42	Source 42
43. Feature 43	Source 43
44. Feature 44	Source 44
45. Feature 45	Source 45
46. Feature 46	Source 46
47. Feature 47	Source 47
48. Feature 48	Source 48
49. Feature 49	Source 49
50. Feature 50	Source 50
51. Feature 51	Source 51
52. Feature 52	Source 52
53. Feature 53	Source 53
54. Feature 54	Source 54
55. Feature 55	Source 55
56. Feature 56	Source 56
57. Feature 57	Source 57
58. Feature 58	Source 58
59. Feature 59	Source 59
60. Feature 60	Source 60
61. Feature 61	Source 61
62. Feature 62	Source 62
63. Feature 63	Source 63
64. Feature 64	Source 64
65. Feature 65	Source 65
66. Feature 66	Source 66
67. Feature 67	Source 67
68. Feature 68	Source 68
69. Feature 69	Source 69
70. Feature 70	Source 70
71. Feature 71	Source 71
72. Feature 72	Source 72
73. Feature 73	Source 73
74. Feature 74	Source 74
75. Feature 75	Source 75
76. Feature 76	Source 76
77. Feature 77	Source 77
78. Feature 78	Source 78
79. Feature 79	Source 79
80. Feature 80	Source 80
81. Feature 81	Source 81
82. Feature 82	Source 82
83. Feature 83	Source 83
84. Feature 84	Source 84
85. Feature 85	Source 85
86. Feature 86	Source 86
87. Feature 87	Source 87
88. Feature 88	Source 88
89. Feature 89	Source 89
90. Feature 90	Source 90
91. Feature 91	Source 91
92. Feature 92	Source 92
93. Feature 93	Source 93
94. Feature 94	Source 94
95. Feature 95	Source 95
96. Feature 96	Source 96
97. Feature 97	Source 97
98. Feature 98	Source 98
99. Feature 99	Source 99
100. Feature 100	Source 100

```

1. .920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cidne="CS0DK011VB20"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/cidne_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	11.6%	Score	716.4	DB	13	Length	920																																																																																																
Best Local Similarity	95.9%	Pred. No.	9e-159																																																																																																				
Matches	746	Conservative	0	Mismatches	31	Indels	1	Gaps	1																																																																																														
QY	5265	CTCTTCCAAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGCAGCGGACGAGGAGGAGGA	5324	DB	66	CTCTTCCAAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGCAGCGGACGAGGAGGA	125	QY	5325	GTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATCTTTACCGAAGACGTAGCCG	5384	DB	126	GTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATCTTTACCGAAGACGTAGCCG	185	QY	5385	CACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGCTTACAGAGAACCTCATCT	5444	DB	186	CACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGCTTACAGAGAACCTCATCT	245	QY	5445	GTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACAGCTCCTCAGCAGG	5504	DB	246	GTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACAGCTCCTCAGCAGG	305	QY	5505	GACCCCTGCGCGAGCGTACCTGGACATCCCGAAACCGCGCTACCTGGCGCCTTGCCATTTC	5564	DB	306	GACCCCTGCGCGAGCGTACCTGGACATCCCGAAACCGCGCTACCTGGCGCCTTGCCATTTC	365	QY	5565	CTCAGGAGCGAATTTACTTTGGCGTCTCATACACAGGATAAATTAAGGGTCATTTCGTGCAA	5624	DB	366	CTCAGGAGCGAATTTACTTTGGCGTCTCATACACAGGATAAATTAAGGGTCATTTCGTGCAA	425	QY	5625	GGGAAACCTCGTGAAGGAGTCCGGCACTGAAACCAACCGGGGCCGCTCACCTCCCGCAG	5684	DB	426	GGGAAACCTCGTGAAGGAGTCCGGCACTGAAACCAACCGGGGCCGCTCACCTCCCGCAG	485	QY	5685	CAGCCCAACAGCGAGGCCACCCACGTATCAAACGAGCACATACCAAGCGCGTGGCGTC	5744	DB	486	CAGCCCAACAGCGAGGCCACCCACGTATCAAACGAGCACATACCAAGCGCGTGGCGTC	545	QY	5745	CAGCCCAAGCGCGCCGAAGGCCCCAGCCACCGCGAGAGCAACACCCCAACCGGTA	5804	DB	546	CAGCCCAAGCGCGCCGCGA - GCCCGAGCCACCCGCGAGAGCAACCCCAACCGGTA	604	QY	5805	CCCGAGGGGCGGACCGAGCTGCGCAGGGAACAAGTCTCTTGCGCGCCCTTGGAGCGGAGA	5864	DB	605	CCCGAGGGGCGGACCGAGCTGCGCAGGGAACAAGTCTCTTGCGCGCCCTTGGAGCGGAGA	664	QY	5865	GAAGTCCCCCGGCGATGCTCAGACCGCGAGAGAGCGGTCCCCCGGAGGCTGTTTGA	5924	DB	665	GAAGTCCCCCGGCGGATGCTCAGACCGCGAGAGAGCGGTCCCCCGGAGGCTGTTTGA	724	QY	5925	AGACAGCAGCAGGGGCCGGCTGCTTGCCTGCGGAGCGGTGAGGACCCCGCTGTGCCAGGTGAA	5984	DB	725	AGACAGCAGCAGGGGCCGGCTGCTTGCCTGCGGAGCGGTGAGGACCCCGCTGTGCCAGGTGAA	784	QY	5985	CAAGGGAAGGGCAGAGTGCTCTCAAGTTTTCACGGTTTAACTGTCACTATTAT	6042	DB	785	CAAGGCTTGGACACAGCTTTCAGTATAAATCTCAGCCAGAAAAACCACTCCTCATCT	842

RESULT 6

RESULT	6
BQ228524	
LOCUS	BQ228524 879 bp mRNA linear EST 02-MAY-2002
DEFINITION	AGENCOURT_7600932 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5061485 5', mRNA sequence.
ACCESSION	BQ228524
VERSION	BQ228524.1 GI:20409924
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 879) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D. Email: cgabbs-x@mail.nih.gov
JOURNAL	Tissue Procurement: ARCC/DCTD/DTF
COMMENT	cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM1331 row: g column: 22 High quality sequence stop: 636.

FEATURES

```

1. .879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6061485"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Skin; Vector: pCMV-SPORT6; Site.1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."

```

ORIGIN

Query Match	11.2%	Score 692	DB 13	Length 879
Best Local Similarity	100.0%	Pred. No. 5.6e-153		
Matches 692	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	3946	CGGAGGAAGCTGCCACCGCAAGCAACGGACACCCACACCCATCCACGCGCAGCCACC	4005	
Db	1	CGGAGGAAGCTGCCACCGCAAGCAACGGACACCCACACCCATCCACGCGCAGCCACC	60	
Qy	4006	GCAGGGCAGCAGATGCGCATGTCGGCCATCGTGCGGTGCAGAGACACAGCCCGAGTGC	4065	
Db	61	GCAGGGCAGCAGATGCGCATGTCGGCCATCGTGCGGTGCAGAGACACAGCCCGAGTGC	120	
Qy	4066	ATGAGCCTGTGCGCCCGCCCATCCAGCGCAGAAAGGAGTCTTCACTCCAGAGGAATTT	4125	
Db	121	ATGAGCCTGTGCGCCCGCCCATCCAGCGCAGAAAGGAGTCTTCACTCCAGAGGAATTT	180	
Qy	4126	AGTCGGCGTCTTAAGGAACGATGCACCAATATTCTCCACGATTCACGTAGGACTG	4185	
Db	181	AGTCGGCGTCTTAAGGAACGATGCACCAATATTCTCCACGATTCACGTAGGACTG	240	
Qy	4186	AACATGGAGGCCACAAAGTGTGTGTGTCTGTGATACCGTGCACACTTGGACGCCAGGCA	4245	
Db	241	AACATGGAGGCCACAAAGTGTGTGTGTCTGTGATACCGTGCACACTTGGACGCCAGGCA	300	
Qy	4246	TCCAAATGTCTCGAATGTCAAGTGATGTGCACCCCAAGTGCTCCAGTGCCTTGGCCAGCC	4305	
Db	301	TCCAAATGTCTCGAATGTCAAGTGATGTGCACCCCAAGTGCTCCAGTGCCTTGGCCAGCC	360	
Qy	4306	ACCTGCGGCTTGCTGTCTGAATATGCCACACACTTCCACCGAGGCCCTTCCCGGTGACAAA	4365	

Db 361 ACCTGGGGCTTGCTGCTGATATGCGACACACTTCACCGAGGCGCTTGTGCGGTGACAAA 420
 Qy 4366 ATGAATCTCCCGAGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGACCTGGAAGGGTGG 4425
 Db 421 ATGAATCTCCCGAGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGACCTGGAAGGGTGG 480
 Qy 4426 ATGAAGTGGCCAGGATACCAACAGGAGACCAAGGCTGGGACAGGAAGTACATTGTC 4485
 Db 481 ATGAAGTGGCCAGGATACCAACAGGAGACCAAGGCTGGGACAGGAAGTACATTGTC 540
 Qy 4486 CTGAGGAGTCAAAAGTCCTCATTTATGACATGAAGCCAGAGAACTGCGACAGAGCGCG 4545
 Db 541 CTGAGGAGTCAAAAGTCCTCATTTATGACATGAAGCCAGAGAACTGCGACAGAGCGCG 600
 Qy 4546 GTGGAAGAAATTTAGCTGTGCTTCCGACGGGGATGTATCTATTCAATGGTCCGTTGTT 4605
 Db 601 GTGGAAGAAATTTAGCTGTGCTTCCGACGGGGATGTATCTATTCAATGGTCCGTTGTT 660
 Qy 4606 GCTTCGGAATCTCGCAATACAGCCAAAGCAGA 4637
 Db 661 GCTTCGGAATCTCGCAATACAGCCAAAGCAGA 692

RESULT 7
 BC031156
 LOCUS 958 bp mRNA linear HTC 04-MAR-2003
 DEFINITION Mus musculus, similar to citron, clone IMAGE:4976752, mRNA.
 ACCESSION BC031156
 VERSION BC031156.1 GI:21411076
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 958)
 Direct Submission
 Submitted (03-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 59 Row: j Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis
 This clone has the following problem: retained intron.

FEATURES
 source
 Location/Qualifiers
 1..958
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4976752"
 /tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH103"
 /note="Vector: pCMV-SPORT6"

ORIGIN
 Query Match 11.2%; Score 688.8; DB 11; Length 958;
 Best Local Similarity 85.4%; Pred. No. 3.5e-152;
 Matches 781; Conservative 0; Mismatches 127; Indels 6; Gaps 1;
 Qy 5149 AATTACAGTATCTTCATTGGAACCAATAAATTTACGAAATCGACATGAAGCAGTACACG 5208
 Db 40 AAGAAACGCAATCTTCATTGGGACCAACAAATTCATGAGATCGACATGAAGCAGTACACG 99
 Qy 5209 CTCGAGGAATTCCTGGGATGAAGATGACATTCTCTTGGCACCTGCTGTGTGTGGCGCTCT 5268
 Db 100 CTTGATGAGTTCCTGGACAGAGACGACCATTCCTTGGCACTGCTGTGTTCGCTCTCTCG 159
 Qy 5269 TCCAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCCAGGCGAGCGAGAGAGTAC 5328
 Db 160 TCCAACAGCTTCCCTGCTCCATTGTGCAGCGCAACAGCGCGGCGAGGAGAAATAC 219
 Qy 5329 TTGCTGTGTTTCCACAGAAATTTGGAGTCTTCGTGATTCTTTACGGAAGACGTAGCCGACA 5388
 Db 220 CTGCTGTGTTTCCACGAAATTTGGGTGTTCGTGATTCTTTACGGAAGACGTAGCCGACA 279
 Qy 5389 GACGATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTTTGCCTTACAGAGAAACCTATCTGTTT 5448
 Db 280 GATGATCTTAAGTGGAGTCTGCTTACCTCTGGCCTTGGCCTTACAGAGAAACCTATCTGTTT 339
 Qy 5449 GTGACCCACTTCACTCACTCGAAGTAATTGAGATCCAGGACGCTCTCCAGCAGGAGCC 5508
 Db 340 GTGACTCACCTTCACTCCCTCTGGGAGTCAITGAGATCCAGGCCAGATCTCTCCTGGGAGC 399
 Qy 5509 CCTGCCGAGCGTACCTGGACATCCCGAACCCGCGGTACCTGGGCCCTGTGCATTTCTCTCA 5568
 Db 400 CTTGCCCGAGCATATCTGGAATTCGAAATTCGAAATTCGCTGCTGCTGCGCCGCGATTTCTCTCC 459
 Qy 5569 GGACGATTTACTTGGGCTCTCTATACAGGATTAATTAAGGTCTATTGCTCTCAAGGA 5628
 Db 460 GGACGATTTACTTGGGCTCTCTATACAGGATTAATTAAGGTCTATTGCTCTCAAGGA 519
 Qy 5629 AACTCTGTGAAGGAGTCCGGCACTGAAACACACCGGGGCCGCTCCACCTCCCGCAGCAGC 5688
 Db 520 AACTCTGTGAAGGAGTCCGGCACTGAGCAGCACCGGGTGGCTCCACCTCCCGCAGCAGC 579
 Qy 5689 CCAACAGCGAGCGCCACCCAGTACACAGGACATACCAAGCGCGTGGCTCTCAGC 5748
 Db 580 CCAACAGCGAGCGCCACCAACATACACAGGACATACCAACAGCGCGTGGCTCTCAGC 639
 Qy 5749 CCAGCGCGCCCGAAGCGCCCGAGCCAGCCAGGAGCCAGGACACCCACCGCTACC-- 5806
 Db 640 CCGCGCGCCCGAAGCGCCCGAGCCAGCCAGGAGCCAGGACACCGGACCGCTACCGA 659
 Qy 5807 ----GCGAGGGCGGACCGGAGCTGCGCAGGGAACAAGTCTCTGCGCGCGCCCTCTGGAGCGA 5862
 Db 700 GACAGAGAGGGTCCGACAGAGCTGCGCAGGGAACAAGTCTCCAGGCGCGCTCTGGAGCGG 759
 Qy 5863 GAGAGTCTCCCGCGCGGATGCTCAGCACCGGAGAGCGGCTCCCGCGGAGGCTGTTT 5922
 Db 760 GAGAGTCTCCCGCGCGGATGCTCAGCACCGGAGAGCGGCTCCCGCGGAGGACTGTTT 819
 Qy 5923 GAGAGACAGCAGCGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCCAGGTG 5982
 Db 820 GAGAGACAGCAGCGGGCGCGCTGCTGCGGAGCGCTGAGGAGCCCGCTGTCCCAGGT 879
 Qy 5983 AACAGGGAAGGCGAGAGTCTCTCAAGTCTTACGGTTTACCGTTTACACTGTCCACTTAT 6042
 Db 880 AACAGGCTCTGGGACCGAGTCTTCAAGTCTTACCGTTTACCGTTTACCGTTTACCGTTT 939
 Qy 6043 GACTGGAATAAAAA 6056
 Db 940 AAAAAAATAAAAAA 953

RESULT 8
 BF905370/c

LOCUS BF905370 652 bp mRNA linear EST 18-JAN-2001
DEFINITION IL3-MT0267-261200-410-H07 MT0267 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF905370
VERSION BF905370.1 GI:12296829
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 652)
AUTHORS Dias Neto,E., Garcia Correa,R., Varjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PubMed 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-MT0267-
261200-410-H07&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 624.
Location/Qualifiers
1..652
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0267"
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 10.2%; Score 629.2; DB 10; Length 652;
Best Local Similarity 98.0%; P-adj. No. 4e-138;
Matches 637; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 2477 GGAACATGAAGCCCAAGAGATGATTTCTGAACCTCGGCAACAGAAATTTTACCTGG 2536
DB 652 GGAACATGAAGCCCAAGAGATGATTTCTGAACCTCGGCAACAGAAATTTTACCTGG 593
QY 2537 AGACACAGCTGGGAAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAGAAGA 2596
DB 592 AGACACAGCTGGGAAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAGAAGA 533
QY 2597 TCAGCCACCAAGACCAAGATGACAGATCGCTGCTGGAACTGGAGCAAGATTGCGGG 2656
DB 532 TCAGCCACCAAGACCAAGATGACAGATCGCTGCTGGAACTGGAGCAAGATTGCGGG 473
QY 2657 AGTCAGCTAGAGCAGGAGGAGCAAGAACTGGAGCTCAAGCGCCAGCTTCACAGAGCTAC 2716
DB 472 AGTCAGCTAGAGCAGGAGGAGCAAGAACTGGAGCTCAAGCGCCAGCTTCACAGAGCTAC 413
QY 2717 AGCTCTCCCTGCAGGAGCGCGAGTCACAGTTGACAGCCTCGAGGCTGCACGGCGGCC 2776

DB 412 AGCTCTCCCTGCAGGAGCGCGAGTCACAGTTGACAGCCTGCAGGCTGCACGGCGGCC 353
QY 2777 TGGAGAGCCAGCTTGGCAGGCGGAAGACAGAGCTGGAAGAGACACAGAGAGAGCTGAAG 2836
DB 352 TGGAGAGCCAGCTTGGCAGGCGGAAGACAGAGCTGGAAGAGACACAGAGAGAGCTGAAG 293
QY 2837 AGGAGATCCAGGCACTCAGCGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCTTTC 2896
DB 292 AGGAGATCCAGGCACTCAGCGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCTTTC 233
QY 2897 GTAACAGCTTACTGTAAATCACAGACTCGGAGAGAGCTTAAACACAGCTGACCGAGACA 2956
DB 232 GTAACAGCTTACTGTAAATCACAGACTCGGAGAGAGCTTAAACACAGCTGACCGAGACA 173
QY 2957 ACCTGGAATCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCG 3016
DB 172 ACCTGGAATCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCG 113
QY 3017 CCAACGACGAGATTTGACAACTCGAAGTGAAGTGAACCTCTCCGCGGAGATCACGG 3076
DB 112 CCAACGACGAGATTTGACAACTCGAAGTGAAGTGAACCTCTCCGCGGAGATCACGG 53
QY 3077 AACGAGATGCGAGCTTACCAGCAGAGCAAAACGATGGAGGCTCTGAAG 3126
DB 52 AACGAGATGCGAGCTTACCAGCAGAGCAAAACGATGGAGGCTCTGAAG 3
RESULT 9
LOCUS BG976452 881 bp mRNA linear EST 12-JUN-2001
DEFINITION 602846269f1 NC1_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4976752 5',
mRNA sequence.
ACCESSION BG976452
VERSION BG976452.1 GI:14364089
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAMi0970 row: j column: 17
High quality sequence stop: 841.
Location/Qualifiers
1..881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4976752"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SmaI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
ORIGIN
Query Match 9.8%; Score 606.6; DB 12; Length 881;


```
Db 301 CAGGTGGTGGAGTTGGCGGTGAAGGAACAACAGGCCGAGATTCTGTGCTTCGACAGGCT 360
QY 3397 CTCRAAGAGCAGAACTCAAGGCCGAGAGCCTCTCTGACAAAGCTCAATGACTCGAGAAG 3456
Db 361 CTCRAAGGAGCAGAACTCAAGGCCGAGAGCCTGTGCGACAAGCTCAAGACCTCGAGAAG 420
QY 3457 AAGCATGCTATGCTTGAATGAATGCCGAGACTTACAGCGAAGCTTACAGCGAAGCTGAGACHTGAACGA 3516
Db 421 AAGCAGCGCATGCTGGAGTGAACGCCCGGAGCTTACAGCAGAACTAGAGACAGAGCGG 480
QY 3517 GAGCTCAAAACAGAGGCTTCTGGAAGCAGCAAGCAAAATTACAGCAGCAGATGGACTCTGCAG 3576
Db 481 GAGCTCAAAACAGAGGCTTCTGGAAGCAGCAAGCAAAATTGAGCAGCAGATGGACTCTGCAG 540
QY 3577 AAAATCAATTTTCGCTGCTGACTCAAGGACTGCAAGAACTCTAGATCGGGCTGACTA 3636
Db 541 AAGAACCAATCTTTCAGACTGACGCAAGGCTGACGAGGCGCTGACCGGCGGATCTG 600
QY 3637 CTGAGACGAGAAAGTGAATGAGGCACTATTTCTCAACAAACCAAACTCAATTTGTTCT 3756
Db 601 CTGAAGACAGAAAGGAGGCACTGAGTACCAAGCTGGAAGAACTCAAGGTTCTCTACTCT 660
QY 3697 CATGAAAAGGTGAATGGAAGGCACTATTTCTCAACAAACCAAACTCAATTTGTTCTG 3756
Db 661 CACGAGAAAGTGAATGGAAGGCACTCTCTCAGCANA-CANACTCAATTTGTTCTG 719
QY 3757 CAGGCCAAATGGACCACTGCTTAAAA 3784
Db 720 CAAG-CANAATGGACCACTGCTTAAAA 746

RESULT 11
LOCUS BG912161
DEFINITION 602812B33F1 NCI_CGAP Brn67 Homo sapiens cDNA clone IMAGE:4944657
5', mRNA sequence.
ACCESSION BG912161
VERSION BG912161.1 GI:14292637
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 843)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10890 row: a column: 10
High quality sequence stop: 778.
FEATURES
Location/Qualifiers
1..843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4944657"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="pH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn67"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
```

```
Query Match 9.5%; Score 584.6; DB 12; Length 843;
Best Local Similarity 92.8%; Pred. No. 2e-127;
Matches 669; Conservative 0; Mismatches 44; Indels 8; Gaps 5;

QY 4636 GAAAAAGCAGAACTCATGCTAAACCTTGGAACCTCCCTGCTGAACTCGGAAGTGAT 4695
Db 122 GAAAAAGCAGAACTCATGCTAAACCTTGGAACCTCCCTGCTGAACTCGGAAGTGAT 181
QY 4696 GACCGTCTAGACATGAATGACGCTGCCCTTCAGTGACCAAGTGCTGTGTGGGCAAC 4755
Db 182 GACCGTCTAGACATGAATGACGCTGCCCTTCAGTGACCAAGTGCTGTGTGGGCAAC 241
QY 4756 GAGGAAGGGCTCTACGCCCTGAATGCTTGAAGAACTCCCTAACCATGCTCCAGGAATT 4815
Db 242 GAGGAAGGGCTCTACGCCCTGAATGCTTGAAGAACTCCCTAACCATGCTCCAGGAATT 301
QY 4816 GGAAGCAGTCTTCCAAATTTATATTATCAAGGACCTCGAAGAGCTACTCATGATAGCAG 4875
Db 302 GGAAGCAGTCTTCCAAATTTATATTATCAAGGACCTCGAAGAGCTACTCATGATAGCAG 361
QY 4876 GAAAGAGGGGCACTGTGCTTTGTGGAAGCTGAAGAAAGTGAACAGTCCCTGGGCCAGTCC 4935
Db 362 GAAAGAGGGGCACTGTGCTTTGTGGAAGCTGAAGAAAGTGAACAGTCCCTGGGCCAGTCC 481
QY 4936 CACCTGCTGCCAGCCGACATCTCACCAACATTTTGAAGCTGTCAAGGGCTGCCAC 4995
Db 422 CACCTGCTGCCAGCCGACATCTCACCAACATTTTGAAGCTGTCAAGGGCTGCCAC 481
QY 4996 TTGTTTTGGGCGAGCAAGATTGAGAAAGGGCTCTGATCTGTGACGCCATGCCAGCAAA 5055
Db 482 TTGTTTTGGGCGAGCAAGATTGAGAAAGGGCTCTGATCTGTGACGCCATGCCAGCAAA 541
QY 5056 GTGCTCATCTCTCGCTACACGAAACCTCAGCAATATCTGCAATCCGGAAGAGATAGAG 5115
Db 542 GTGCTCATCTCTCGCTACACGAAACCTCAGCAATATCTGCAATCCGGAAGAGATAGAG 601
QY 5116 ACCTCAGAGCCCTGCAGCTGTATCCACTTCCCAATTAACAGTATCTCTCAATT-GGAACCAA 5174
Db 602 ACCTCAGAGCCCTGCAGCTGTATCCACTTCCCAATTAACAGTATCTCTCAATTGGGAACCAA 661
QY 5175 TAAATCTTAGC-AAATCGACATGAA-GCAGTACACCTCGAGGAATT--CCTGGATAAGA 5230
Db 662 TAAATCTTAGCAAAATCGACATGAAGGCAGTACACGCTCGAGGAATTCTCTGGATAACGA 721
QY 5231 ATGACCATTCCTTGGCACCTGCTGTG---TTTCCGCCCTCTTCCAAACAGCTTCCCTGTCT 5287
Db 722 ATGACCATTCCTTGGCACCTGCTGTG---TTTCCGCCCTCTTCCAAACAGCTTCTCTCTC 781
QY 5288 CAATCCTGCAAGGTGAACAGCGCAGGGCAGCAGAGAGTACTTGTGTGTTTCCAGGAAT 5347
Db 782 AATCGTGGCAGGTGAACAGGGCACGCGCAGGAGAGAGTATTGCTGTGTTCCCGGACT 841
QY 5348 T 5348
Db 842 T 842

RESULT 12
LOCUS CF744580
DEFINITION UI-M-GVO-clt-n-23-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
IMAGE:30617710 5', mRNA sequence.
ACCESSION CF744580
VERSION CF744580.1 GI:37640920
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 647)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
(This clone was contributed by the Brain Molecular Anatomy Project (BMAP))

Seq primer: pyX-5.

FEATURES

source
Location/Qualifiers
1..647
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30617710"
/tissue_type="whole brain"
/dev_stage="1.5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
/note="Organ: Brain; Vector: pyX-Asc; Site: 1: EcoR I; Site: 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACGGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 9.3%; Score 573.6; DB 14; Length 647;
Best Local Similarity 93.8%; Pred. No. 6.9e-125;
Matches 608; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
QY 1131 CACCCTCAAGTCGACGATGACACCTCCAAATTTGATGAACAGAGAAGAAATTCGTGGGT 1190
DB 1 CACCCTCAAGTCGACGATGACACCTCCAAATTTGATGAACAGAGAAGAAATTCGTGGGT 60
QY 1191 TTCATCTCTCCGTGCGACCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGT 1250
DB 61 TTCATCTCTGTGTGCCAGCTGAGCCCTCGGCTTCTCAGCGAAGAGCTGCCGTTTGT 120
QY 1251 GGGGTTTTCGTACAGCAAGGCACTGGGATCTTGGTAGATCTGAGTCTGTGTGTCGGG 1310
DB 121 GGGATTTTCGTACAGCAAGGCACTGGGATCTTGGTAGATCTGAGTCTGTGTGTCGAG 180
QY 1311 TCTGGACTCCCTGCCAAGCTAGCTCCATGGAAGAAATCTTCATCAAAAGCAAGA 1370
DB 181 TCTGGACTCCCTGCCAAGCTAGCTCCATGGAAGAAATCTTCATCAAAAGCAAGA 240
QY 1371 GCTACAAGACTCTCAGGACAAGTGTCAAGAATGGAGCAGGAATAATGACCCGGTTACATCG 1430
DB 241 GCTCCAAGACTCCAGGACAAGTGTCAAGAATGGAGCAGGAATAATGACCCGGTTACATCG 300
QY 1431 GAGAGTGTGAGAGGTGGAGCTGCTTAGTCAGAGGAGTGGAGCTGAAGGCTCTGA 1490
DB 301 CAGAGTGTGAGAGGTGGAGCTGCTTAGTCAGAGGAGTGGAGCTGAAGGCTCTGA 360
QY 1491 GACTCAGAGATCCCTCTCGAGCAGGACCTTCTACCTACATCAAGAAATGCAAGTGT 1550
DB 361 GACTCAGAGATCCCTCTCGAGCAGGACCTTCTACCTACATCAAGAAATGCAAGTGT 420
QY 1551 AAAGCGAGTTTGGAGCAAGCAGCTGAGGTGTCCAGGAGGATGACAAGCACTGCA 1610

Db 421 AAAGCGAGTTTGGAGCAAGCGGATGGAGGTGCCAGAGGATGACAAGCTCTGCA 480
QY 1611 GCTTCTCCATGATATCAGAGCAGAGCCGAGAGCTCCAGAAATCAAGAGCAGAGTA 1670
DB 481 GCTTCTCCAGACATCCGAGCAGAGCCGAGAGCTCCAGAGATCAAGAGCAGAGTA 540
QY 1671 CCAGGCTCAAGTGAAGAAATGAGTTGATGATGAATCAGTTCGAAAGAGGATCTTGTC 1730
DB 541 CCAGGCTCAGTGGAGAGATGAGGCTGATGATGATCAGTTCGAAAGAGGATCTTGTC 599
QY 1731 AGCAAGAGCAGGAGTGTCTTACGAATCTGAGTTCGAGAGAGTCTCG 1778
DB 600 AGCCCGCAGACCGAGCGATCTCTACGAGTCTGAGCTGAGGAGTCTCG 647

RESULT 13

CF723360 730 bp mRNA linear EST 09-OCT-2003
LOCUS UI-M-GVO-cjh-j-20-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
DEFINITION IMAGE:30546187 5', mRNA sequence.
ACCESSION CF723360
VERSION CF723360.1 GI:37597528
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 730)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyX-5.

FEATURES

source
Location/Qualifiers
1..730
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30546187"
/tissue_type="whole brain"
/dev_stage="1.5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
/note="Organ: Brain; Vector: pyX-Asc; Site: 1: EcoR I; Site: 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACGGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 9.3%; Score 572.8; DB 14; Length 730;
Best Local Similarity 88.7%; Pred. No. 1.2e-124;

RESULT 15
 BO807302
 LOCUS
 DEFINITION
 NISC_kk01b10.y1 NCI_CGAP_Brn72 Macaca mulatta cDNA clone
 IMAGE:5330250 5', mRNA sequence.
 ACCESSION
 BO807302
 VERSION
 BO807302.1 GI:22031511
 EST.
 KEYWORDS
 Macaca mulatta (rhesus monkey)
 SOURCE
 ORGANISM
 Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.
 REFERENCE
 1 (bases 1 to 640)
 AUTHORS
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL
 Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 CDNA Library Preparation:
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 info@image.llnl.gov
 Plate: LLAM11838 row: C column: 19
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1..640
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /db_xref="taxon:9544"
 /clone="IMAGE:5330250"
 /tissue_type="hypothalamus"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP_Brn72"
 /note="Organ: Brain; Vector: pCMV-SPORT6.1; Site 1: NotI;
 Site 2: EcoRV; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.2 Kb. Constructed by Invitrogen.
 Note: this is a NCI_CGAP Library."

FEATURES

source

ORIGIN

Query Match 9.0%; Score 556.8; DB 13; Length 640;
 Best Local Similarity 98.8%; Pred. No. 6.9e-121;
 Matches 561; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 4070 GCCTGTGCCCCCGCATCCAGCCGACGAGAGGAGTCTTCACTCCAGAGGAATTAGTC 4129
 Db 1 GCCTGTGCCCCCGCATCCAGCCGACGAGAGGAGTCTTCACTCCAGAGGAATTAGTC 60
 QY 4130 GCGCTTTAAGGAACGCGATGACCAACAATATTCCTCACCGATTCAACGTAGGACTGAACA 4189
 Db 61 GCGCTTTAAGGAACGCGATGACCAACAATATTCCTCACCGATTCAACGTAGGACTGAACA 120
 QY 4190 TCGGAGCCACAAGTGTGCTGTCTGGATACCGTGCACCTTTGGACGCCAGGATCCA 4249
 Db 121 TCGGAGCCACAAGTGTGCTGTCTGGATACCGTGCACCTTTGGACGCCAGGATCCA 180
 QY 4250 AATGTCTCGAATGTCTGAGTGTCTACCCCAAGTGTCCACGTGCTTGCAGCCACCT 4309
 Db 181 AATGTCTCGAATGTCTGAGTGTCTACCCCAAGTGTCCACGTGCTTGCAGCCACCT 240
 QY 4310 GCGGCTTGGCTCTGAATATGCGCACACACTTCAACGAGGCTTCTGCGTGACAAAATGA 4369
 Db 241 GCGGCTTGGCTCTGAATATGCGCACACACTTCAACGAGGCTTCTGCGTGACAAAATGA 300
 QY 4370 ACTCCCCAGGCTCTCAGACCAAGGAGCCAGCAGAGCTTGCACCTGGAAGGTTGGATGA 4429
 Db 301 ACTCCCCAGGCTCTCAGACCAAGGAGCCAGCAGAGCTTGCACCTGGAAGGTTGGATGA 360
 QY 4430 AGGTGCCCGAGGAATAACAAACAGGACAGCAAGGCTGGGACAGGAAGTACATTCTCTCTGG 4489

Search completed: July 3, 2004, 19:29:05
 Job time : 9474 secs

Db 361 AGGTGCCCGAGGAATAACAAACAGGACAGCAAGCTGGGACAGGAAGTACATTCTCTGG 420
 QY 4490 AGGGATCAAAAAGTCTCTCATTTTATGCAATGAAGCCAGAGAAGCTGGACAGAGGCCGGTGG 4549
 Db 421 AGGGATCAAAAAGTCTCTCATTTTATGCAATGAAGCCAGAGAAGCTGGACAGAGGCCGGTGG 480
 QY 4550 AAGAATTTGAGCTGTGCTTCCCGACCGGGATGTATCTATTTCATGTGCGGTTGGTGCCTT 4609
 Db 481 AAGAATTTGAGCTGTGCTTCCCGACCGGGATGTATCTATTTCATGTGCGGTTGGTGCCTT 540
 QY 4610 CCGAACTCGCAATAACAGCCAAAGCAGA 4637
 Db 541 CCGAACTCGCAATAACAGCCAAAGCAGA 568

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 16:52:20 ; Search time 1715 Seconds
(without alignments)
17324.953 Million cell updates/sec

Title: US-10-017-216-3
Perfect score: 6159
Sequence: 1 atgttgaagtccaatatgg.....ttctgagaacagattatgc 6159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6159	100.0	6159	14	US-10-017-216-3
2	6159	100.0	6162	15	US-10-325-430-11
3	6159	100.0	6574	14	US-10-017-216-1
4	6159	100.0	6574	15	US-10-325-430-10
5	5666.4	92.0	6298	13	US-10-415-011-43
6	5661.4	91.9	6165	14	US-10-028-946-1
7	5631	91.4	6189	13	US-10-262-511-1
8	5629.8	91.4	6189	13	US-09-964-956-10
9	5629.8	91.4	6201	13	US-10-262-511-13
10	5629.8	91.4	6201	13	US-09-964-956-8
11	5358	87.0	5877	14	US-10-028-946-3
12	2564.8	41.6	3131	13	US-10-276-774-137
13	2414	39.2	2542	13	US-10-262-511-7
14	2317.2	37.6	2497	13	US-10-262-511-5

15	1841	29.9	1870	13	US-10-262-511-3	Sequence 3, Appli
16	1524.2	24.7	1870	13	US-10-262-511-9	Sequence 9, Appli
17	1467.6	23.8	1915	13	US-10-262-511-11	Sequence 11, Appl
18	1393.4	22.6	2066	17	US-10-311-034-33	Sequence 33, Appl
19	1391.8	22.6	1515	9	US-09-804-471A-1	Sequence 1, Appli
20	1391.8	22.6	1515	15	US-10-238-709-1	Sequence 1, Appli
21	1391.8	22.6	1515	17	US-10-724-594-1	Sequence 1, Appli
22	952.4	15.5	2162	16	US-10-120-988-419	Sequence 419, App
23	950.6	15.4	957	16	US-10-412-897-1	Sequence 1, Appli
24	750	12.2	1133	15	US-10-282-048-1	Sequence 1, Appli
25	701.4	11.4	995	13	US-10-276-774-255	Sequence 255, App
26	421.4	6.8	2870	13	US-10-423-114-26241	Sequence 26241, A
27	350.4	5.7	354	14	US-10-040-739-9	Sequence 9, Appli
28	258	4.2	258	16	US-10-305-720-513	Sequence 513, App
29	243.6	4.0	1530	13	US-10-415-011-24	Sequence 24, Appl
30	243.6	4.0	4698	17	US-10-702-496-1	Sequence 1, Appli
31	243.6	4.0	5373	13	US-10-333-314-40	Sequence 40, Appl
32	236.8	3.8	6335	16	US-10-388-934-85	Sequence 85, Appl
33	234	3.8	5373	17	US-10-702-496-5	Sequence 5, Appli
34	228.2	3.7	2785	13	US-10-342-887-930	Sequence 930, App
35	228.2	3.7	2785	13	US-10-172-118-930	Sequence 26, Appl
36	228.2	3.7	5694	13	US-10-362-892-26	Sequence 26, Appl
37	228.2	3.7	5694	16	US-10-288-798-26	Sequence 39, Appl
38	228.2	3.7	7151	17	US-10-433-794-39	Sequence 179, App
39	225	3.7	2621	15	US-10-252-157-179	Sequence 99, Appl
40	221.6	3.6	5347	16	US-10-210-130-99	Sequence 101, App
41	221.6	3.6	5875	16	US-10-210-130-101	Sequence 1, Appli
42	220.8	3.6	3407	9	US-09-971-845-1	Sequence 9, Appli
43	220.8	3.6	3407	12	US-10-380-235-9	Sequence 1449, Ap
44	208.4	3.4	3363	16	US-10-012-697-1449	Sequence 9, Appli
45	208.4	3.4	6780	12	US-10-399-225-9	

ALIGNMENTS

RESULT 1
US-10-017-216-3
; Sequence 3, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPILLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-216-3

Query Match	100.0%;	Score	6159;	DB	14;	Length	6159;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	6159;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATCGTGGTGTGCTGAACCCATT	60				
Db	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATCGTGGTGTGCTGAACCCATT	60				
Qy	61	GCCAGCGGGCTCCAGGCTGATCTCTTTCAGGGGAAACACCCCTTATGACTCAA	120				
Db	61	GCCAGCGGGCTCCAGGCTGATCTCTTTCAGGGGAAACACCCCTTATGACTCAA	120				
Qy	121	CAGCAGATGCTCTCTCTTTCAGGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA	180				
Db	121	CAGCAGATGCTCTCTCTTTCAGGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA	180				

181 QY GAATGCAGTCAGCCTGCTCTGATGAAGATTAAACACGTGAGCAACTTTTGTCCGGAAAGTAT 240
181 Db GAATGCAGTCAGCCTGCTCTGATGAAGATTAAACACGTGAGCAACTTTTGTCCGGAAAGTAT 240
241 QY TCCGACACCATAGCTAGCTGAGTTTACAGGAGCTCCAGCCTTCGCAAAAGGACTTCGAAGTCAGA 300
241 Db TCCGACACCATAGCTAGCTGAGTTTACAGGAGCTCCAGCCTTCGCAAAAGGACTTCGAAGTCAGA 300
301 QY AGTCTTGAGGTGTTGGTGCACCTTTGCTGAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG 360
301 Db AGTCTTGAGGTGTTGGTGCACCTTTGCTGAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG 360
361 QY GACATCTATCTATGAAAGTGATGAAGAGAGAGGCTTTATTTGGCCAGGAGCAGGTTTCA 420
361 Db GACATCTATGCTATGAAGTGATGAAGAGAGAGGCTTTATTTGGCCAGGAGCAGGTTTCA 420
421 QY TTTTGTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480
421 Db TTTTGTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480
481 QY CAGTATGCTCTTTCAGGACCAAAATCACTTTATCTGATGAGGAATATCAGCCTGGAGGG 540
481 Db CAGTATGCTCTTTCAGGACCAAAATCACTTTATCTGATGAGGAATATCAGCCTGGAGGG 540
541 QY GACTTGTCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATGAATCAAGTTT 600
541 Db GACTTGTCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATGAATCAAGTTT 600
601 QY TACTAGCTAGCTGATTTTGGCTGTTTCAAGGTTTCACTGATGGGATACGTGATCGA 660
601 Db TACTAGCTAGCTGATTTTGGCTGTTTCAAGGTTTCACTGATGGGATACGTGATCGA 660
661 QY GACATCAAGCTCAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTTGATTTT 720
661 Db GACATCAAGCTCAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTTGATTTT 720
721 QY GGATCTGCCGCAAAATGAATTCGTTGACCGCACAGGAGTGAATGCCAAACTCCCGATTGGACC 780
721 Db GGATCTGCCGCAAAATGAATTCGTTGACCGCACAGGAGTGAATGCCAAACTCCCGATTGGACC 780
781 QY CCAGATTACATGCTCCTCAAGTGTCTGACTGTGATCAACGGGATGGAAGGACCTTAC 840
781 Db CCAGATTACATGCTCCTCAAGTGTCTGACTGTGATCAACGGGATGGAAGGACCTTAC 840
841 QY GGCCTGGAGTGTGACTGGTGGTCAAGTGGCGTGAATGCCCTATGAGATGATTTATGGAGA 900
841 Db GGCCTGGAGTGTGACTGGTGGTCAAGTGGCGTGAATGCCCTATGAGATGATTTATGGAGA 900
901 QY TCCGCCCTCGCAGAGGACCTTCCAGAACCTTCAATTAACATTTATGAAATTCAGCGG 960
901 Db TCCGCCCTCGCAGAGGACCTTCCAGAACCTTCAATTAACATTTATGAAATTCAGCGG 960
961 QY TTTTGTGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTTGATCTGATTCAAAGC 1020
961 Db TTTTGTGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTTGATCTGATTCAAAGC 1020
1021 QY TTGTTGTGGCCGAGAAAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTCTTTCTTC 1080
1021 Db TTGTTGTGGCCGAGAAAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTCTTTCTTC 1080
1081 QY TCTAAATTTGACTTGAACCAACTTGTAACTCTCTCCCTCCCTTCTGTTCCACCTCAAG 1140
1081 Db TCTAAATTTGACTTGAACCAACTTGTAACTCTCTCCCTCCCTTCTGTTCCACCTCAAG 1140
1141 QY TCTGAGATGACCTTCAATTTTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1200
1141 Db TCTGAGATGACCTTCAATTTTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1200
1201 QY CGTGCAGCTCAGCCCTCAGGCTTCTCGGGTGAAGACTGCGGTTTGTGGGTTTTCG 1260
1201 Db CGTGCAGCTCAGCCCTCAGGCTTCTCGGGTGAAGACTGCGGTTTGTGGGTTTTCG 1260
1261 QY TACAGCAAGGCACCTGGGATTTCTTGGTAGATCTGTAGTCTGTGTGTGGGCTCGGACTCC 1320

1261 Db TACAGCAAGGCACCTGGGATTTCTTGGTAGATCTGTGTGTGGGCTCGGACTCC 1320
1321 QY CCTGCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAGCAAGAGACTACAAGAC 1380
1321 Db CCTGCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAGCAAGAGACTACAAGAC 1380
1381 QY TCTCAGACAAGTGTCTCAAGATGGAGAGAAATGACCCGGTTACATCCGAGAGTGTCA 1440
1381 Db TCTCAGACAAGTGTCTCAAGATGGAGAGAAATGACCCGGTTACATCCGAGAGTGTCA 1440
1441 QY GAGTGTGAGGCTGTGCTTAGTCAGAACGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
1441 Db GAGTGTGAGGCTGTGCTTAGTCAGAACGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
1501 QY TCCCTCTCGAGCAGGACCTTGTCTTACATACAGAAATGACAGTGTAAAGCGAAGT 1560
1501 Db TCCCTCTCGAGCAGGACCTTGTCTTACATACAGAAATGACAGTGTAAAGCGAAGT 1560
1561 QY TTGGAGCAAGCAGGATGGAGGTGTCCTCAGAGAGATGACAAAGCAGCTGAGCTTCTCCAT 1620
1561 Db TTGGAGCAAGCAGGATGGAGGTGTCCTCAGAGAGATGACAAAGCAGCTGAGCTTCTCCAT 1620
1621 QY GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTAGCAGGCTCAA 1680
1621 Db GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTAGCAGGCTCAA 1680
1681 QY GTGGAAGAAATGAGTTGATGATGATGATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740
1681 Db GTGGAAGAAATGAGTTGATGATGATGATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740
1741 QY CGGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800
1741 Db CGGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800
1801 QY CGGAAGCGACAGAAATGTCAGCTTAACTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAA 1860
1801 Db CGGAAGCGACAGAAATGTCAGCTTAACTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAA 1860
1861 QY GTGGAGAAATCGAAATCGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
1861 Db GTGGAGAAATCGGAGAAATCGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
1921 QY CTCAAGAGAAAATCGGAGAGGCTGCAAAAGGAGCGAGAGGAGGCTGGAAGAGCTG 1980
1921 Db CTCAAGAGAAAATCGGAGAGGCTGCAAAAGGAGCGAGAGGAGGCTGGAAGAGCTG 1980
1981 QY CAGAACCGAGAGATTTCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTGAGGACGC 2040
1981 Db CAGAACCGAGAGATTTCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTGAGGACGC 2040
2041 QY CGCATTTCTCTGGAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGA 2100
2041 Db CGCATTTCTCTGGAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGA 2100
2101 QY CTGAAGATGACATCCAGACAAATCCCAAGCTAGAGAGATGCTGATTAAGTCTG 2160
2101 Db CTGAAGATGACATCCAGACAAATCCCAAGCTAGAGAGATGCTGATTAAGTCTG 2160
2161 QY GAGCTCGAAGAGAAAATCGGAGGCGCCCAAGTCTCAGCCAGCAGCCTAGAGTGCACCTG 2220
2161 Db GAGCTCGAAGAGAAAATCGGAGGCGCCCAAGTCTCAGCCAGCAGCCTAGAGTGCACCTG 2220
2221 QY AAACAGAAAGAGCAGCACTATGAGGAAAAGATTAAAGTGTGGACAAATCAGATAAGAAA 2280
2221 Db AAACAGAAAGAGCAGCACTATGAGGAAAAGATTAAAGTGTGGACAAATCAGATAAGAAA 2280
2281 QY GACTTGTGACAAAGAGACACTGGAAGAAATGATGATCAGAGACACGAGGAGGAGGCCAT 2340
2281 Db GACTTGTGACAAAGAGACACTGGAAGAAATGATGATCAGAGACACGAGGAGGAGGCCAT 2340
2341 QY GAGAAAGGCAAAATCTCAGCGCAACAGAGGCGATGATCAATGCTATGATTTCCAAAGTCC 2400

Db 2341 GAGAGGGCAAAATTTCTCAGCGAACAGAGGGCGATGATCAATGCTATGGATTCCAAGATC 2400
QY 2401 AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAATAAATCTTGCAGCAAAATAGC 2460
Db 2401 AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAATAAATCTTGCAGCAAAATAGC 2460
QY 2461 AGTCCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTTGAATCTCAGGCCAA 2520
Db 2461 AGTCCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTTGAATCTCAGGCCAA 2520
QY 2521 CAGAAATTTTACTCTGGAGACACAGGCTGGAGTTGGAGGCCCGAGAACCGGAAACTGGAG 2580
Db 2521 CAGAAATTTTACTCTGGAGACACAGGCTGGAGTTGGAGGCCCGAGAACCGGAAACTGGAG 2580
QY 2581 GAGCAGCTGGAGAAATCAGCCACCAAGACACACAGTGCACAGAAATCGGCTGCTGGAATG 2640
Db 2581 GAGCAGCTGGAGAAATCAGCCACCAAGACACACAGTGCACAGAAATCGGCTGCTGGAATG 2640
QY 2641 GAGCAAGATTCGGGAGGTCAGTCTAGAGCACGAGGACGAGAACTGGAGCTCAAGGCG 2700
Db 2641 GAGCAAGATTCGGGAGGTCAGTCTAGAGCACGAGGACGAGAACTGGAGCTCAAGGCG 2700
QY 2701 CAGCTCACAGAGCTACAGCTCTCCCTGCAAGAGCGGAGTCAAGTTGACAGCCCTGGAG 2760
Db 2701 CAGCTCACAGAGCTACAGCTCTCCCTGCAAGAGCGGAGTCAAGTTGACAGCCCTGGAG 2760
QY 2761 GCTGCACGCGGCGCCCTGGAGAGCCAGCTTCCGACAGGCGGAGACAGAGCTGGAAGAGACC 2820
Db 2761 GCTGCACGCGGCGCCCTGGAGAGCCAGCTTCCGACAGGCGGAGACAGAGCTGGAAGAGACC 2820
QY 2821 ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAAATCCAGGCG 2880
Db 2821 ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAAATCCAGGCG 2880
QY 2881 AAATTTGATGCTTTCTGTAACAGCTGTACTGTAAATCACAGACCTGGAGAGCAGCTAAGC 2940
Db 2881 AAATTTGATGCTTTCTGTAACAGCTGTACTGTAAATCACAGACCTGGAGAGCAGCTAAGC 2940
QY 2941 CAGCTGACCGAGGACAAACGCTGAACTCAACAAACCAAAACTTTACTTTCTCCAAACAACTC 3000
Db 2941 CAGCTGACCGAGGACAAACGCTGAACTCAACAAACCAAAACTTTACTTTCTCCAAACAACTC 3000
QY 3001 GATGAGGCTTTGGGCGCCAAACAGCAGATTTGACAACTGCGNAGTGAAGTGCACCATCTC 3060
Db 3001 GATGAGGCTTTGGGCGCCAAACAGCAGATTTGACAACTGCGNAGTGAAGTGCACCATCTC 3060
QY 3061 CGCCGGAGATCACGGAACGAGAGATGACAGTTACAGCCAGAACAAACGATGAGGCT 3120
Db 3061 CGCCGGAGATCACGGAACGAGAGATGACAGTTACAGCCAGAACAAACGATGAGGCT 3120
QY 3121 CTGAAGACCACTGTCACCATGCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTAAACGAT 3180
Db 3121 CTGAAGACCACTGTCACCATGCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTAAACGAT 3180
QY 3181 GAGCTGTAGAAAAGAGCGGCACTGGAGGCGCTCGGAGAGGCTCGGCTGATGAGAAA 3240
Db 3181 GAGCTGTAGAAAAGAGCGGCACTGGAGGCGCTCGGAGAGGCTCGGCTGATGAGAAA 3240
QY 3241 TCCCAAGTTTGAAGTGTGCGGTTCCAGAGCTGCAGAGAAATGCTGGACACCGAGAAAACAGAGC 3300
Db 3241 TCCCAAGTTTGAAGTGTGCGGTTCCAGAGCTGCAGAGAAATGCTGGACACCGAGAAAACAGAGC 3300
QY 3301 AGGGCAGAGCCGATCAGCGGATTCACCGAGTCTCGCCAGGTTGGAGTGGAGTGGAGTGAAG 3360
Db 3301 AGGGCAGAGCCGATCAGCGGATTCACCGAGTCTCGCCAGGTTGGAGTGGAGTGGAGTGAAG 3360
QY 3361 GAGCAACAGGCTGAGATTTCTGCTCTGACAGAGGCTTCTCAAAGACAGAGCTGAAGGCC 3420
Db 3361 GAGCAACAGGCTGAGATTTCTGCTCTGACAGAGGCTTCTCAAAGACAGAGCTGAAGGCC 3420
QY 3421 GAGAGCCTCTCTGACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3480
Db 3421 GAGAGCCTCTCTGACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3480

QY 3481 GCCCGAAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540
Db 3481 GCCCGAAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540
QY 3541 GAGCAAGCCAAATTTACAGCAGCAGATGGACCTGCAGAAAATCACAATTTTCCGTCTGACT 3600
Db 3541 GAGCAAGCCAAATTTACAGCAGCAGATGGACCTGCAGAAAATCACAATTTTCCGTCTGACT 3600
QY 3601 CAAGGACTCAAGAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAAGTGAAGTTCG 3660
Db 3601 CAAGGACTCAAGAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAAGTGAAGTTCG 3660
QY 3661 GAGTATCAGCTGAAAACATTTCAAGTTCTCTATTCTCATGAAAAGTGAATAATGGAAGGC 3720
Db 3661 GAGTATCAGCTGAAAACATTTCAAGTTCTCTATTCTCATGAAAAGTGAATAATGGAAGGC 3720
QY 3721 ACTATTTCTCAACAAACCAAACTCATTTCTGCAAGCCAAAATGGAACCACTGCT 3780
Db 3721 ACTATTTCTCAACAAACCAAACTCATTTCTGCAAGCCAAAATGGAACCACTGCT 3780
QY 3781 AAAAAGAAAAGGTTTATTTAGTTCGACGGAAGAGACCCCTGCTTTTACCCACACAGGTT 3840
Db 3781 AAAAAGAAAAGGTTTATTTAGTTCGACGGAAGAGACCCCTGCTTTTACCCACACAGGTT 3840
QY 3841 CCTCTGCACTACAAATGAGCTGAAAGCTGGCCCTCGAAGAGGAGAAAGCTGCTGCAAG 3900
Db 3841 CCTCTGCACTACAAATGAGCTGAAAGCTGGCCCTCGAAGAGGAGAAAGCTGCTGCAAG 3900
QY 3901 CTGAGGAAGCCCTTCAAGAGACCCGATCAGCTCGGCTCGCCCGGGAGAGGCTGCC 3960
Db 3901 CTGAGGAAGCCCTTCAAGAGACCCGATCAGCTCGGCTCGCCCGGGAGAGGCTGCC 3960
QY 3961 CACGCAAAACCAACGAGACCCACACCCATCCACGCGACCCACCGGAGGAGCAGATC 4020
Db 3961 CACGCAAAACCAACGAGACCCACACCCATCCACGCGACCCACCGGAGGAGCAGATC 4020
QY 4021 GCCATGTCGCCATTTCTGCGGTGCCAGACACAGCCCACTGATGACCTGCTGCGCC 4080
Db 4021 GCCATGTCGCCATTTCTGCGGTGCCAGACACAGCCCACTGATGACCTGCTGCGCC 4080
QY 4081 CCGCCATCCAGCGCAGAAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTCCGCGCTTTAAG 4140
Db 4081 CCGCCATCCAGCGCAGAAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTCCGCGCTTTAAG 4140
QY 4141 GAACGCAATGCACCAATATTTCTTCAACCGATTCACAGTAGGACTGAACATGCGAGCCACA 4200
Db 4141 GAACGCAATGCACCAATATTTCTTCAACCGATTCACAGTAGGACTGAACATGCGAGCCACA 4200
QY 4201 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGAGCCAGGCATCCAAATGCTCTGAA 4260
Db 4201 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGAGCCAGGCATCCAAATGCTCTGAA 4260
QY 4261 TGTGAGTGTGTGTCAACCCAAAGTGTCTTCCAGCTGCTTGCAGCCACTCGCGGTTGCCT 4320
Db 4261 TGTGAGTGTGTGTCAACCCAAAGTGTCTTCCAGCTGCTTGCAGCCACTCGCGGTTGCCT 4320
QY 4321 GCTGAATATGCCACACACTTCAACCGAGCTTCTGCGGTGACAAAATGAATGAACTCCCGAGGT 4380
Db 4321 GCTGAATATGCCACACACTTCAACCGAGCTTCTGCGGTGACAAAATGAATGAACTCCCGAGGT 4380
QY 4381 CTCAGACCAAGGAGCCAGCAGCTTGCACCTGGAAGGGTGGATGAAGTGCAGG 4440
Db 4381 CTCAGACCAAGGAGCCAGCAGCTTGCACCTGGAAGGGTGGATGAAGTGCAGG 4440
QY 4441 AATAACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAA 4500
Db 4441 AATAACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAA 4500
QY 4501 GTCCTCATTTATGAACAATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAGAAATTTGAG 4560
Db 4501 GTCCTCATTTATGAACAATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAGAAATTTGAG 4560

4561 CTGTGCTTCCGACGGGATGTATCTATTCTATGTCGGTTCCTTCGAACTCGCA 4620
4561 CTGTGCTTCCGACGGGATGTATCTATTCTATGTCGGTTCCTTCGAACTCGCA 4620
4621 AATACAGCCAAAGCAGAAAAGCAGAAAGCTGATGCTTAAACTGTCTGGAAACTCCCTGCTG 4680
4621 AATACAGCCAAAGCAGAAAAGCAGAAAGCTGATGCTTAAACTGTCTGGAAACTCCCTGCTG 4680
4681 AACTGGAAGTGATGACCGCTTAGACATGAATGACAGCTGCCCTTCAAGTACACAGGTG 4740
4681 AACTGGAAGTGATGACCGCTTAGACATGAATGACAGCTGCCCTTCAAGTACACAGGTG 4740
4741 GTGTGTGGTGGCCACGAGGAAGGGCTCTACGCCCTGAATGTCTTGAAGAACTCCCTTAACC 4800
4741 GTGTGTGGTGGCCACGAGGAAGGGCTCTACGCCCTGAATGTCTTGAAGAACTCCCTTAACC 4800
4801 CATGTCCAGGAATTTGAGAGAGCTTTCCAAATTTATATTATTAAGGACCTTGAGAGCTTA 4860
4801 CATGTCCAGGAATTTGAGAGAGCTTTCCAAATTTATATTATTAAGGACCTTGAGAGCTTA 4860
4861 CTGATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGAGTGAAGAAAGTGAACAG 4920
4861 CTGATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGAGTGAAGAAAGTGAACAG 4920
4921 TCCTGGCCAGTCCCACTTCCTGCCCAGCCGACATCTCACCCAACTTTTGAAGCT 4980
4921 TCCTGGCCAGTCCCACTTCCTGCCCAGCCGACATCTCACCCAACTTTTGAAGCT 4980
4981 GTCAAGGGTGCACACTGTTTGGGGCAGGCAAGATTGAGAAAGGGCTGTGATCTGTGCA 5040
4981 GTCAAGGGTGCACACTGTTTGGGGCAGGCAAGATTGAGAAAGGGCTGTGATCTGTGCA 5040
5041 GCCATGCCAGCAAAGTCTGCTCATCTCCGCTACACGAAACCTCAGCAAATCTGCATC 5100
5041 GCCATGCCAGCAAAGTCTGCTCATCTCCGCTACACGAAACCTCAGCAAATCTGCATC 5100
5101 CGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCACTTCCAAATTTACAGTATC 5160
5101 CGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCACTTCCAAATTTACAGTATC 5160
5161 CTCATTGGAAACCAATAAATTTCTACGAAATCGCATGAAGCAGTACACGCTCGAGGAATTC 5220
5161 CTCATTGGAAACCAATAAATTTCTACGAAATCGCATGAAGCAGTACACGCTCGAGGAATTC 5220
5221 CTGGATAAGAAATGACCAATTCCTTGGCACCTGCTGTGTTTGGCCCTCTTCCAAACAGCTTC 5280
5221 CTGGATAAGAAATGACCAATTCCTTGGCACCTGCTGTGTTTGGCCCTCTTCCAAACAGCTTC 5280
5281 CCTGCTCAATCGTGAGGTACACGGCAGGCGCAGGAGGAGTACTTGTGTGTTTC 5340
5281 CCTGCTCAATCGTGAGGTACACGGCAGGCGCAGGAGGAGTACTTGTGTGTTTC 5340
5341 CAGGAATTTGGAGTGTCTGGATTCTTACGGAAGACGTAGCCGACAGACGATCTCAAG 5400
5341 CAGGAATTTGGAGTGTCTGGATTCTTACGGAAGACGTAGCCGACAGACGATCTCAAG 5400
5401 TGGAGTCGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTATCTGTTGTGACCCACTTC 5460
5401 TGGAGTCGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTATCTGTTGTGACCCACTTC 5460
5461 AACTCACTCGAAGTAATGAGATCCAGGACCGCTCCTCAGCAGGAGCCCTCCCGGAGCG 5520
5461 AACTCACTCGAAGTAATGAGATCCAGGACCGCTCCTCAGCAGGAGCCCTCCCGGAGCG 5520
5521 TACTGGACATCCGAAACCCGCGCTACCTGGGCCCTGCAATTTCTCAGGAGCGAATTTAC 5580
5521 TACTGGACATCCGAAACCCGCGCTACCTGGGCCCTGCAATTTCTCAGGAGCGAATTTAC 5580
5581 TTGGCGTCTCTATACAGAGATAAATTAAGGGTCATTGCTGCAAGGAAACCTCGTGAAG 5640
5581 TTGGCGTCTCTATACAGAGATAAATTAAGGGTCATTGCTGCAAGGAAACCTCGTGAAG 5640
5641 GAGTCCGGCACTGAACACCAACCGGGCCGCTCCACTCCCGCAGCAGCCCAACAGCGA 5700

Db 5641 GAGTCCGGCACTGAACACACCGGGGCCGCTCCACTCCCGCAGCAGCCCAACAGCGA 5700
Qy 5701 GGCCCAACCCAGTAAACAGGACACATCAACAGCGCGTGGCTCCAGCCAGCGCCGCC 5760
Db 5701 GGCCCAACCCAGTAAACAGGACACATCAACAGCGCGTGGCTCCAGCCAGCGCCGCC 5760
Qy 5761 GAAGGCCCCAGCCACCGCGAGAGCGCAAGCACACCCACCGCTACCCGAGGGCGGACC 5820
Db 5761 GAAGGCCCCAGCCACCGCGAGAGCGCAAGCACACCCACCGCTACCCGAGGGCGGACC 5820
Qy 5821 GAGTGTGGCGAGGACAAAGTCTCTGTGGCCGCCCTCTGAGCGAGAGAGTCCCCCGGCCG 5880
Db 5821 GAGTGTGGCGAGGACAAAGTCTCTGTGGCCGCCCTCTGAGCGAGAGAGTCCCCCGGCCG 5880
Qy 5881 ATGCTCAGCAGCGGGAGAGAGCGGTCCCCCGGGAGGCTTTTGAAGACACAGCAGGGGC 5940
Db 5881 ATGCTCAGCAGCGGGAGAGAGCGGTCCCCCGGGAGGCTTTTGAAGACACAGCAGGGGC 5940
Qy 5941 CGGCTGCTCGGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAGGGAAGAGGGCAG 6000
Db 5941 CGGCTGCTCGGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAGGGAAGAGGGCAG 6000
Qy 6001 AGTCCCTCTCAAGTTTTCACGGTTAACACTGTCACTTATGACTGGAATAAAAAAGCTG 6060
Db 6001 AGTCCCTCTCAAGTTTTCACGGTTAACACTGTCACTTATGACTGGAATAAAAAAGCTG 6060
Qy 6061 GACAACTGCCAGTAACTGTCAGTCCCTGAGGATCATCCAGCTGAATGAGAAATCCGG 6120
Db 6061 GACAACTGCCAGTAACTGTCAGTCCCTGAGGATCATCCAGCTGAATGAGAAATCCGG 6120
Qy 6121 CAGCAGGTGAAAAGTCTGTTCTGAGAACAGATTATTC 6159
Db 6121 CAGCAGGTGAAAAGTCTGTTCTGAGAACAGATTATTC 6159

RESULT 2
US-10-325-430-11
; Sequence 11, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL D-SORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838,336 and 52908
; FILE REFERENCE: MPI01-294P1RNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 6162
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6162)
US-10-325-430-11

Query Match 100.0%; Score 6159; DB 15; Length 6162;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTGTAAGTTCAAATATGAGCGCGGAATCTTTGATGCTGCTGCTGAACCAATT 60
Db 1 ATGTGTAAGTTCAAATATGAGCGCGGAATCTTTGATGCTGCTGCTGAACCAATT 60
Qy 61 GCACGCCGGCCTCCAGGCTGAATCTGTTCTCAGGGNAACCACTTTATGACTCAA 120

Db 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
Qy 121 CAGCAGATGTCCTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 121 CAGCAGATGTCCTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Qy 181 GAATCAGTCAGCTGCTCTGATGAAATTAAGACAGCTGAGCAACTTTGTCGGGAAGTAT 240
Db 181 GAATCAGTCAGCTGCTCTGATGAAGATTAAGCACTGAGCAACTTTGTCGGGAAGTAT 240
Qy 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 300
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 300
Qy 301 AGCTCTGAGTGTGCTCACTTTCTGTAAGTCAGGTGCTAAGAGAGAAAGCAACCGGG 360
Db 301 AGCTCTGAGTGTGCTCACTTTCTGTAAGTCAGGTGCTAAGAGAGAAAGCAACCGGG 360
Qy 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCACGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCACGAGCAGGTTTCA 420
Qy 421 TTTTGTGAGGAGCGGAAACATATTATCTGAAGCACAAGCCGCTGGATCCCCAATTA 480
Db 421 TTTTGTGAGGAGCGGAAACATATTATCTGAAGCACAAGCCGCTGGATCCCCAATTA 480
Qy 481 CAGTATGCTTTACAGACAAAATCACCTTTATCTGATGGAGGATATCAGCCTGGAGG 540
Db 481 CAGTATGCTTTACAGACAAAATCACCTTTATCTGATGGAGGATATCAGCCTGGAGG 540
Qy 541 GACTTGTCTCACTTTTGAATAGATGAGGACCAAGTTAGATGAATACTGATACAGTTT 600
Db 541 GACTTGTCTCACTTTTGAATAGATGAGGACCAAGTTAGATGAATACTGATACAGTTT 600
Qy 601 TACCTAGCTGACTGATTTGGCTGTTACACGGTTCTATGATGGGATAGCTGATCGA 660
Db 601 TACCTAGCTGACTGATTTGGCTGTTACACGGTTCTATGATGGGATAGCTGATCGA 660
Qy 661 GACATCAAGCTTGAGAACATTTCTGTTGACCCGACAGGACATCAAGCTGGTGAATTT 720
Db 661 GACATCAAGCTTGAGAACATTTCTGTTGACCCGACAGGACATCAAGCTGGTGAATTT 720
Qy 721 GGATCTGCCGCGAAATGAATTCAAACAGATGTTGAATGCCAACTCCCGATTGGGACC 780
Db 721 GGATCTGCCGCGAAATGAATTCAAACAGATGTTGAATGCCAACTCCCGATTGGGACC 780
Qy 781 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGACCTTAC 840
Db 781 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGACCTTAC 840
Qy 841 GGCCTGGACTGTGACTGGTGGTCACTGGGCGTGAATTCATGAGATGATTTATGGGAGA 900
Db 841 GGCCTGGACTGTGACTGGTGGTCACTGGGCGTGAATTCATGAGATGATTTATGGGAGA 900
Qy 901 TCCCTCTTCGAGAGGGAACCTTCGCGAGAACCTTCAATAAGATTATGAATTCAGCGG 960
Db 901 TCCCTCTTCGAGAGGGAACCTTCGCGAGAACCTTCAATAAGATTATGAATTCAGCGG 960
Qy 961 TTTTTCGAATTTCCAGATGACCCCAAGTGAGCAGTGAATTTCTGATCTGATTCAAAGC 1020
Db 961 TTTTTCGAATTTCCAGATGACCCCAAGTGAGCAGTGAATTTCTGATCTGATTCAAAGC 1020
Qy 1021 TTGTTGTGGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGGCAATCTTTCTTC 1080
Db 1021 TTGTTGTGGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGGCAATCTTTCTTC 1080
Qy 1081 TCTAAAATTTGACTGGAAACAACTTCGTAATCTCTCCCTCCCTTCGTTCCACCTCAAG 1140
Db 1081 TCTAAAATTTGACTGGAAACAACTTCGTAATCTCTCCCTCCCTTCGTTCCACCTCAAG 1140
Qy 1141 TCTGACGATGACACCTTCCAAATTTGATGAACCAAGAAATTCGTGGGTTTCATCTCTCT 1200
Db 1141 TCTGACGATGACACCTTCCAAATTTGATGAACCAAGAAATTCGTGGGTTTCATCTCTCT 1200

Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCGCTTTGTGGGTTTTTCG 1260
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCGCTTTGTGGGTTTTTCG 1260
Qy 1261 TACAGCAAGSCACTGGGATCTTTGTAGATCTGAGTCTGTTGTTGTCGGGCTCTGGACTCC 1320
Db 1261 TACAGCAAGSCACTGGGATCTTTGTAGATCTGAGTCTGTTGTTGTCGGGCTCTGGACTCC 1320
Qy 1321 CTTGCCAAGACTAGCTCCATTTGAAAAAAACTTTCTCATCAAAAAGCAAGCTTACAGAC 1380
Db 1321 CTTGCCAAGACTAGCTCCATTTGAAAAAAACTTTCTCATCAAAAAGCAAGCTTACAGAC 1380
Qy 1381 TCTCAGCAAGTGTCAAGATGGAGCGAATGACCCGGTTACATCGAAGTGTCA 1440
Db 1381 TCTCAGCAAGTGTCAAGATGGAGCGAATGACCCGGTTACATCGAAGTGTCA 1440
Qy 1441 GAGGTGAGGCTCTGCTTTAGTCAGAAAGGAGTGAAGCCCTCTGAGACTCAGAGA 1500
Db 1441 GAGGTGAGGCTCTGCTTTAGTCAGAAAGGAGTGAAGCCCTCTGAGACTCAGAGA 1500
Qy 1501 TCCCTCTGAGCAGGACCTTGTCTACCTACATCAGAAATGCAAGTGTAAAGCGAAGT 1560
Db 1501 TCCCTCTGAGCAGGACCTTGTCTACCTACATCAGAAATGCAAGTGTAAAGCGAAGT 1560
Qy 1561 TTGAGCAAGCAGGATGGAGGTGCCAGAGGATCAAAAAGCACTGCAAGCTTCTCCAT 1620
Db 1561 TTGAGCAAGCAGGATGGAGGTGCCAGAGGATCAAAAAGCACTGCAAGCTTCTCCAT 1620
Qy 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAAGAGCAGGATACCAAGGCTCAA 1680
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAAGAGCAGGATACCAAGGCTCAA 1680
Qy 1681 GTGGAAGAAATGAGTTGATGATCACTTGGAGAGGATCTTGTCTCAGCAAGAAGA 1740
Db 1681 GTGGAAGAAATGAGTTGATGATCACTTGGAGAGGATCTTGTCTCAGCAAGAAGA 1740
Qy 1741 CCGAGTCTCTACGAATCTGAGCTGAGAGACTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Db 1741 CCGAGTCTCTACGAATCTGAGCTGAGAGACTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Qy 1801 CGAAAGCGACAGAACTGAGCACTTGAAGCTTAAAGCTAAGGATCAAGGGAAGCCCTGAA 1860
Db 1801 CGAAAGCGACAGAACTGAGCACTTGAAGCTTAAAGCTAAGGATCAAGGGAAGCCCTGAA 1860
Qy 1861 GTGGGAATATGCAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1861 GTGGGAATATGCAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Qy 1921 CTCGAGAGAACTGGAGAGGCTGCAAAAGGAGCGAGCCGAGAGGAGCTGGAGAACTG 1980
Db 1921 CTCGAGAGAACTGGAGAGGCTGCAAAAGGAGCGAGCCGAGAGGAGCTGGAGAACTG 1980
Qy 1981 CAGAACCGAGAGGATTTCTTGAAGGATCAGAAAGAGTGTGTGAGGCTGAGAACTG 2040
Db 1981 CAGAACCGAGAGGATTTCTTGAAGGATCAGAAAGAGTGTGTGAGGCTGAGAACTG 2040
Qy 2041 CGCCATTTCTTGGAGAACTAGAGACTAGAGCCATGAGAGCTTGAAGAAACAGA 2100
Db 2041 CGCCATTTCTTGGAGAACTAGAGACTAGAGCCATGAGAGCTTGAAGAAACAGA 2100
Qy 2101 CTGAGGATGACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAATTCG 2160
Db 2101 CTGAGGATGACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAATTCG 2160
Qy 2161 GAGCTCGAAGAAACATCGGAGGCTTCTCAGCCAGCACTTAGAAGTGCACCTG 2220
Db 2161 GAGCTCGAAGAAACATCGGAGGCTTCTCAGCCAGCACTTAGAAGTGCACCTG 2220
Qy 2221 AAAAGAGAGCAGCTATGAGGAAAGATTTAAAGTGTGGCAATCAGATAAAGAAA 2280
Db 2221 AAAAGAGAGCAGCTATGAGGAAAGATTTAAAGTGTGGCAATCAGATAAAGAAA 2280

Qy	2281	GACCTGGCTGACAAGGAGACA	CTCGAGAA	CATGATG	CAGAGACA	CGAGGAGGAGGCCCAT	2340
Db	2281	GACCTGGCTGACAAGGAGACA	CTCGAGAA	CATGATG	CAGAGACA	CACGAGGAGGAGGCCCAT	2340
Qy	2341	GAGAAGGGCAAAATTTCTCAGCGAA	CAGAGCGGATGATCAATG	CTATGGAATTC	CAAGATC	2400	
Db	2341	GAGAAGGGCAAAATTTCTCAGCGAA	CAGAGCGGATGATCAATG	CTATGGAATTC	CAAGATC	2400	
Qy	2401	AGATCCCTTGGAAACAGAGGATTTGGA	ACTGTGAAGCCAAATAA	ACTTTCAGCAAAATAGC	2460		
Db	2401	AGATCCCTTGGAAACAGAGGATTTGGA	ACTGTGAAGCCAAATAA	ACTTTCAGCAAAATAGC	2460		
Qy	2461	AGTCCTTTTACCCAAAGGACATCAAG	CGCCCAAGAACAGATGATTTCTGA	AGCTCAGCGCA	2520		
Db	2461	AGTCCTTTTACCCAAAGGACATCAAG	CGCCCAAGAACAGATGATTTCTGA	AGCTCAGCGCA	2520		
Qy	2521	CAGAAATTTTACCTTGGAGACACAGC	CTGGGAAGTTTGAGGCCCAAGAAC	AGATGATTTCTGAACCTCAGCGCA	2580		
Db	2521	CAGAAATTTTACCTTGGAGACACAGC	CTGGGAAGTTTGAGGCCCAAGAAC	AGATGATTTCTGAACCTCAGCGCA	2580		
Qy	2581	GAGCAGCTGGAGAGATCAGCCACCA	AGACCAAGTGCACAAAGAAATCGG	CTGCTGGAACCTG	2640		
Db	2581	GAGCAGCTGGAGAGATCAGCCACCA	AGACCAAGTGCACAAAGAAATCGG	CTGCTGGAACCTG	2640		
Qy	2641	GAGACAAGATTGCGGGAGGTCAGTCT	AGAGCAGGAGCAGAAATCGGAGCT	CAAGCGC	2700		
Db	2641	GAGACAAGATTGCGGGAGGTCAGTCT	AGAGCAGGAGCAGAAATCGGAGCT	CAAGCGC	2700		
Qy	2701	CAGCTCA	CAGAGTACAGCTCTCCCTCGAGAG	CGGAGTCA	CAGTTGACAGCCCTGCAAG	2760	
Db	2701	CAGCTCA	CAGAGTACAGCTCTCCCTCGAGAG	CGGAGTCA	CAGTTGACAGCCCTGCAAG	2760	
Qy	2761	GCTCCACGGCGGCCCTTGGAGAGCC	AGCTTCGCCAGCGCAAGACAGAGCT	TGGAAGAGACC	2820		
Db	2761	GCTCCACGGCGGCCCTTGGAGAGCC	AGCTTCGCCAGCGCAAGACAGAGCT	TGGAAGAGACC	2820		
Qy	2821	ACAGCAGAAAGCTGAAGAGGAGATCC	AGGCACTCAGCGCACATAGATGAAT	TCCAGCGC	2880		
Db	2821	ACAGCAGAAAGCTGAAGAGGAGATCC	AGGCACTCAGCGCACATAGATGAAT	TCCAGCGC	2880		
Qy	2881	AAATTTGATGCTCTTCGTACAGCTGT	ACTGTATACAGACCTCGAGAGAGCT	GAAGCTTAAC	2940		
Db	2881	AAATTTGATGCTCTTCGTACAGCTGT	ACTGTATACAGACCTCGAGAGAGCT	GAAGCTTAAC	2940		
Qy	2941	CAGCTGACCGGAGCAACCGTGA	ACTCAAACCAAAACTTCTACTTGT	TGCCAAACAACTC	3000		
Db	2941	CAGCTGACCGGAGCAACCGTGA	ACTCAAACCAAAACTTCTACTTGT	TGCCAAACAACTC	3000		
Qy	3001	GATGAGGCTTTCTGGCGCAACGAGAG	ATTGTACAACCTCGGAAGTGAAGTGA	AGCTTCAAGTGAACCTC	3060		
Db	3001	GATGAGGCTTTCTGGCGCAACGAGAG	ATTGTACAACCTCGGAAGTGAAGTGA	AGCTTCAAGTGAACCTC	3060		
Qy	3061	CGCCGGGAGATCACGGAACGAGAG	ATGCAGCTTACCAGCCAGAGAACAG	ATGAGGCT	3120		
Db	3061	CGCCGGGAGATCACGGAACGAGAG	ATGCAGCTTACCAGCCAGAGAACAG	ATGAGGCT	3120		
Qy	3121	CTGAAGACACGTGCACCACTGCTG	GAAGAACAGGTCAATGGAATTTG	AGGCCCTTAACGAT	3180		
Db	3121	CTGAAGACACACGTGCACCACTGCTG	GAAGAACAGGTCAATGGAATTTG	AGGCCCTTAACGAT	3180		
Qy	3181	GAGCTGCTCAAAAAGAGCGGCA	GTGGAGGCTCGAGAGCGCTCTCG	GTGATGAGAA	3240		
Db	3181	GAGCTGCTCAAAAAGAGCGGCA	GTGGAGGCTCGAGAGCGCTCTCG	GTGATGAGAA	3240		
Qy	3241	TCCAGCTTTGAGTGTCCGGTTTC	GAGAGCTCGAGAAATGCTCG	AACCGAGAAACAGAGC	3300		
Db	3241	TCCAGCTTTGAGTGTCCGGTTTC	GAGAGAGCTCGAGAAATGCTCG	AACCGAGAAACAGAGC	3300		
Qy	3301	AGGCGGAGAGCCCATCAGCGGA	TACCGAGCTCTCCGAGGTGGT	GGAGCTGCGATGAAG	3360		
Db	3301	AGGCGGAGAGCCCATCAGCGGA	TACCGAGCTCTCCGAGGTGGT	GGAGCTGCGATGAAG	3360		
Qy	3361	GAGCA	CAAGGCTGAGATTTCTCGCTCT	CTGCAAGGCTCTCA	AAGAGCAGAAAGCTGAAGGCC	3420	

[illegible]

```
Db 4441 AATAACAAACGAGCAGCAAGGCTGGGACAGGAAGTACATTGTCTCTGGAGGGATCAAAA 4500
Qy 4501 GTCCTCATTTATGACAAATGAAGCCAGAGAGAGCTGACAGAGGCGGTGGAGAAATTTGAG 4560
Db 4501 GTCCTCATTTATGACAAATGAAGCCAGAGAGAGCTGACAGAGGCGGTGGAGAAATTTGAG 4560
Qy 4561 CTGTGCTCTCCGACGGGGATGTAATCTAATTCATGTCGCGTGGTGGTCTTCCGAATCGCA 4620
Db 4561 CTGTGCTCTCCGACGGGGATGTAATCTAATTCATGTCGCGTGGTGGTCTTCCGAATCGCA 4620
Qy 4621 AATACAGCCAAAGCAGAAAACGAGAGCTGATGCTAACTGCTTGGAAATCTCCCTGCTG 4680
Db 4621 AATACAGCCAAAGCAGAAAACGAGAGCTGATGCTAACTGCTTGGAAATCTCCCTGCTG 4680
Qy 4681 AAACCTGGAAGTGTATGACCGCTCTAGACATGAACCTGACGCTGCCCTTCAGTGACAGGTG 4740
Db 4681 AAACCTGGAAGTGTATGACCGCTCTAGACATGAACCTGACGCTGCCCTTCAGTGACAGGTG 4740
Qy 4741 GTGTGTTGGTGACCGAGAGGCTCTACGCCCTGAATGTCCTTGAATACTCCCTAAC 4800
Db 4741 GTGTGTTGGTGACCGAGAGGCTCTACGCCCTGAATGTCCTTGAATACTCCCTAAC 4800
Qy 4801 CATGTCCCAAGAAATGGAGCAGCTCTTCCAAATTTATATTATCAAGGACCTGGAGAGCTA 4860
Db 4801 CATGTCCCAAGAAATGGAGCAGCTCTTCCAAATTTATATTATCAAGGACCTGGAGAGCTA 4860
Qy 4861 CTCATGATAGCAGGAGAGAGCGGCACTGTGCTCTTGTGGAGCTGGAAGAGTGAACAG 4920
Db 4861 CTCATGATAGCAGGAGAGAGCGGCACTGTGCTCTTGTGGAGCTGGAAGAGTGAACAG 4920
Qy 4921 TCCCTGGCCGATGCCACCTGCTGCCAGCCGACATCTACCCAAATTTTGAAGCT 4980
Db 4921 TCCCTGGCCGATGCCACCTGCTGCCAGCCGACATCTACCCAAATTTTGAAGCT 4980
Qy 4981 GTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAACGGCTCTGCAATCTGTGCA 5040
Db 4981 GTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAACGGCTCTGCAATCTGTGCA 5040
Qy 5041 GCCATGCCAGCAAGTGTGCTATCTCCGCTACAGAGAACTCAGCAATACTGCAATC 5100
Db 5041 GCCATGCCAGCAAGTGTGCTATCTCCGCTACAGAGAACTCAGCAATACTGCAATC 5100
Qy 5101 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCACAAATACAGTATC 5160
Db 5101 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCACAAATACAGTATC 5160
Qy 5161 CTCATTGGAACCAATAAATCTACGAATTCGAAATCGACATGAAGCAGTACACGCTCAGGAATC 5220
Db 5161 CTCATTGGAACCAATAAATCTACGAATTCGAAATCGACATGAAGCAGTACACGCTCAGGAATC 5220
Qy 5221 CTGGATAAGATGACCAATTCCTTGGCAGCTGCTGTTGTTGGCCCTCTTCCACAGCTTC 5280
Db 5221 CTGGATAAGATGACCAATTCCTTGGCAGCTGCTGTTGTTGGCCCTCTTCCACAGCTTC 5280
Qy 5281 CTTGCTCTCAATCGTCAGGTGAACAGCGAGGCGACGAGAGAGTACTGCTGTGTTTC 5340
Db 5281 CTTGCTCTCAATCGTCAGGTGAACAGCGAGGCGACGAGAGAGTACTGCTGTGTTTC 5340
Qy 5341 CACGAATTTGGAGTGTCTGGAATCTTACGGAAGACGTAGCGGACACAGCATCTCAAG 5400
Db 5341 CACGAATTTGGAGTGTCTGGAATCTTACGGAAGACGTAGCGGACACAGCATCTCAAG 5400
Qy 5401 TGGAGTCTGCTTACCTTTGGCCTTTGCTCTACAGAAACCCCTATCTGTTGTGACCCACTTC 5460
Db 5401 TGGAGTCTGCTTACCTTTGGCCTTTGCTCTACAGAAACCCCTATCTGTTGTGACCCACTTC 5460
Qy 5461 AACTCACTCGAAGTAATGAGATCCAGGACCGCTCTCAGAGGGACCCCTGCCGAGCG 5520
Db 5461 AACTCACTCGAAGTAATGAGATCCAGGACCGCTCTCAGAGGGACCCCTGCCGAGCG 5520
Qy 5521 TACCTGGACATCCGAGCCGCTACCTGGGCCCTGCCATTTCTTCAGAGGATTTAC 5580
Db 5521 TACCTGGACATCCCGAACCCGCTACCTGGGCCCTGCCATTTCTTCAGAGGATTTAC 5580
```

```
Qy 5581 TTGGCGTCTCTATACCAGGATAAATTAAGGGTCATTTGCTGCAAGGAAACCTCGTGAAG 5640
Db 5581 TTGGCGTCTCTATACCAGGATAAATTAAGGGTCATTTGCTGCAAGGAAACCTCGTGAAG 5640
Qy 5641 GAGTCCGGCACTGAACACACCGGGCCGTCCACCTCCGCGAGCAGCCCAACAAGCGA 5700
Db 5641 GAGTCCGGCACTGAACACACCGGGCCGTCCACCTCCGCGAGCAGCCCAACAAGCGA 5700
Qy 5701 GGCCCAACCCACGTACACACGAGCAGCATCACCAAGCGGTGGCTCCAGCCAGCGCCGCC 5760
Db 5701 GGCCCAACCCACGTACACACGAGCAGCATCACCAAGCGGTGGCTCCAGCCAGCGCCGCC 5760
Qy 5761 GAAGGCCCCAGCCACCGCGAGAGCCAGACACACCCACCGCTTACCGGAGGGGCGGACC 5820
Db 5761 GAAGGCCCCAGCCACCGCGAGAGCCAGACACACCCACCGCTTACCGGAGGGGCGGACC 5820
Qy 5821 GAGCTGGCAGGAGCAAGTCTCTGCGCGCCCTTGAGCGAGAGAGTCCCGCGCGCG 5880
Db 5821 GAGCTGGCAGGAGCAAGTCTCTGCGCGCCCTTGAGCGAGAGAGTCCCGCGCGCG 5880
Qy 5881 ATGCTACGACCGGAGAGAGCGGTCCCGGAGAGCTGTTTGAAGACAGCAGCAGGGGC 5940
Db 5881 ATGCTACGACCGGAGAGAGCGGTCCCGGAGAGCTGTTTGAAGACAGCAGCAGGGGC 5940
Qy 5941 CGGCTGCTCGCGGAGCGGTGAGGACCCCGCTGTCCAGGTGAACAGGAAAGAGGCGAG 6000
Db 5941 CGGCTGCTCGCGGAGCGGTGAGGACCCCGCTGTCCAGGTGAACAGGAAAGAGGCGAG 6000
Qy 6001 AGTGCCTCTCAAGTTTTTACGGTTTACACACTGTCACTTATGATCGAATAAAGCTG 6060
Db 6001 AGTGCCTCTCAAGTTTTTACGGTTTACACACTGTCACTTATGATCGAATAAAGCTG 6060
Qy 6061 GACAACTCCAGCTAACTGTCAGTCTCAGGATCATCCAGCTGAATGGAGAAATCCGG 6120
Db 6061 GACAACTCCAGCTAACTGTCAGTCTCAGGATCATCCAGCTGAATGGAGAAATCCGG 6120
Qy 6121 CAGCAGGTTGAAAGTCTGTTCTCAGAACAGATTATTGC 6159
Db 6121 CAGCAGGTTGAAAGTCTGTTCTCAGAACAGATTATTGC 6159

RESULT 3
US-10-017-216-1
; Sequence 1, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPLELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote
; FILE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-216-1

Query Match 100.0%; Score 6159; DB 14; Length 6574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATCTGCTGCTGCTGACCCATT 60
Db 19 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATCTGCTGCTGCTGACCCATT 78
Qy 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCACCTTTATGACTCAA 120
```


Db 79 GCCAGCGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAACCAACCTTTATGACTCAA 138
Qy 121 CACGAGTGTCTCTCTCTTTCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 139 CAGCAGATGTCTCTCTTTCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198
Qy 181 GAATGCACTCAGCTGTCTCTGATGAAGATTAAAGCAGTGAAGCACTTTGTCGGAAGTAT 240
Db 199 GAATGCACTCAGCTGTCTCTGATGAAGATTAAAGCAGTGAAGCACTTTGTCGGAAGTAT 258
Qy 241 TCGACACCAATAGTGAAGTACAGAGCTCCAGCTTTCCGCAAGGACTTCGAAGTCAGA 300
Db 259 TCGACACCAATAGTGAAGTACAGAGCTCCAGCTTTCCGCAAGGACTTCGAAGTCAGA 318
Qy 301 AGTCTTGTAGTGTGTGTCTTGTCTGAGTGTGATGAGAGAGAGCAACCCGGG 360
Db 319 AGTCTTGTAGTGTGTGTCTTGTCTGAGTGTGATGAGAGAGAGCAACCCGGG 378
Qy 361 GACATCTATGCTATGAAGTGAAGAGAGAGAGGCTTTATTTGCCCCAGGAGGTTTCA 420
Db 379 GACATCTATGCTATGAAGTGAAGAGAGAGAGGCTTTATTTGCCCCAGGAGGTTTCA 438
Qy 421 TTTTGTGAGAGAGCGGACATATTATCTCGAGCACAGCCCGTGGATCCCCCAATTA 480
Db 439 TTTTGTGAGAGAGCGGACATATTATCTCGAGCACAGCCCGTGGATCCCCCAATTA 498
Qy 481 CAGTATGCCCTTTCAGGACAAATATCACTTTATCTGATGAGGAAATATCAGCTCGAGGG 540
Db 499 CAGTATGCCCTTTCAGGACAAATATCACTTTATCTGATGAGGAAATATCAGCTCGAGGG 558
Qy 541 GACTTGTCTCACTTTTGAATAGATAGGACCAAGTATAGTGAAGAAACCTGATACAGTT 600
Db 559 GACTTGTCTCACTTTTGAATAGATAGGACCAAGTATAGTGAAGAAACCTGATACAGTT 618
Qy 601 TACTAGCTGAGCTGATTTTGGCTGTTCACAGGGTTTCACTGATGGGATACGTGATCGA 660
Db 619 TACTAGCTGAGCTGATTTTGGCTGTTCACAGGGTTTCACTGATGGGATACGTGATCGA 678
Qy 661 GACATCAAGCTGAGAACATTTCTGTTGACCGGACAGGACATCAAGCTGTGATTTT 720
Db 679 GACATCAAGCTGAGAACATTTCTGTTGACCGGACAGGACATCAAGCTGTGATTTT 738
Qy 721 GGATCTGCCGGAATGAATCAAAAGAGTGTGAATGCCAACTCCCGATTCGGACC 780
Db 739 GGATCTGCCGGAATGAATCAAAAGAGTGTGAATGCCAACTCCCGATTCGGACC 798
Qy 781 CCAGATTATACGCTCTCTGAAGTGTGACTGTGATGAACCGGGATCGAAAGGACCTTAC 840
Db 799 CCAGATTATACGCTCTCTGAAGTGTGACTGTGATGAACCGGGATCGAAAGGACCTTAC 858
Qy 841 GGCTCGGATGTGACTGTGTTGCTCAGTGGCGTGTATGCTATGAGTGAATTTATGGAGA 900
Db 859 GGCTCGGATGTGACTGTGTTGCTCAGTGGCGTGTATGCTATGAGTGAATTTATGGAGA 918
Qy 901 TCCCTCTTCGAGAGGAACTCTCGCAGAACCTTCAATAACATTAATGAATTTCCAGGG 960
Db 919 TCCCTCTTCGAGAGGAACTCTCGCAGAACCTTCAATAACATTAATGAATTTCCAGGG 978
Qy 961 TTTTGAATTTCCAGATGACCCCAAGTGAGAGTGAATTTCTTGATCTGATTTCAAAGC 1020
Db 979 TTTTGAATTTCCAGATGACCCCAAGTGAGAGTGAATTTCTTGATCTGATTTCAAAGC 1038
Qy 1021 TTCTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGGTTTGTGCTGCCATCCTTTCTTC 1080
Db 1039 TTCTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGGTTTGTGCTGCCATCCTTTCTTC 1098
Qy 1081 TCTAAATTTGACTGGAAACAATTTGTTAATCTCTCTCCCTCTCGTTCCCAACCTCAAG 1140
Db 1099 TCTAAATTTGACTGGAAACAATTTGTTAATCTCTCTCCCTCTCGTTCCCAACCTCAAG 1158
Qy 1141 TCTGAGATGACACCTTCCAAATTTTGAATGAACAGAGAGAAATTCGTGGGTTTCACTCTCT 1200
Db 1159 TCTGAGATGACACCTTCCAAATTTTGAATGAACAGAGAGAAATTCGTGGGTTTCACTCTCT 1218

Qy 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1260
Db 1219 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1278
Qy 1261 TACAGCAAGCACTGGGATTTCTTGTAGATCTCAGTCTGTGTGTGGGTCTCGACTCC 1320
Db 1279 TACAGCAAGCACTGGGATTTCTTGTAGATCTCAGTCTGTGTGTGGGTCTCGACTCC 1338
Qy 1321 CCGTCCAGCACTAGCTCCATGGAAGAAATCTTCTATCAAAAGCAAGAGCTCAAGAC 1380
Db 1339 CCGTCCAGCACTAGCTCCATGGAAGAAATCTTCTATCAAAAGCAAGAGCTCAAGAC 1398
Qy 1381 TCTCAGCAAACTGTCAAGATGAGCAAGAAATGACCCGGTTTACATCGAGAGTGTCA 1440
Db 1399 TCTCAGCAAACTGTCAAGATGAGCAAGAAATGACCCGGTTTACATCGAGAGTGTCA 1458
Qy 1441 GAGTGGAGGCTGTCTTGTAGTCAAGAGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1459 GAGTGGAGGCTGTCTTGTAGTCAAGAGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518
Qy 1501 TCCCTCTCGAGCAGGACCTTTGCTACCTACATCAAGAAATGCAAGTAAAGCGAAGT 1560
Db 1519 TCCCTCTCGAGCAGGACCTTTGCTACCTACATCAAGAAATGCAAGTAAAGCGAAGT 1578
Qy 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGAGAGATGACAAAGCACTGCAAGTCTTCCAT 1620
Db 1579 TTGAGCAAGCAGGATGGAGGTGTCCAGAGAGATGACAAAGCACTGCAAGTCTTCCAT 1638
Qy 1621 GATATCAGAGCAGAGCGGAGCTCCAGAAATCAAGAGCAGAGTACCGAGCTCAA 1680
Db 1639 GATATCAGAGCAGAGCGGAGCTCCAGAAATCAAGAGCAGAGTACCGAGCTCAA 1698
Qy 1681 GTGGAAGAAATGAGTTGATGATCAATCACTGGAAGAGATCTTGTCTCAGCAAGAAGA 1740
Db 1699 GTGGAAGAAATGAGTTGATGATCAATCACTGGAAGAGATCTTGTCTCAGCAAGAAGA 1758
Qy 1741 CGGAGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTCTGCTGAGAAATTCAG 1800
Db 1759 CGGAGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTCTGCTGAGAAATTCAG 1818
Qy 1801 CGGAAACGACAGAAATGTCAGCATAACTTGAAGCTAAGGATCAAGGAAAGCTGAA 1860
Db 1819 CGGAAACGACAGAAATGTCAGCATAACTTGAAGCTAAGGATCAAGGAAAGCTGAA 1878
Qy 1861 GTGGAGAAATATGCAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
Db 1879 GTGGAGAAATATGCAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938
Qy 1921 CTCCAGAGAACTGGAGAGGCTGCAAAAGCAGCGAGAGGAGCTTGGAGAGAGCTG 1980
Db 1939 CTCCAGAGAACTGGAGAGGCTGCAAAAGCAGCGAGAGGAGCTTGGAGAGAGCTG 1998
Qy 1981 CAGAACCGAGAGGATTTCTTCTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGAAAGC 2040
Db 1999 CAGAACCGAGAGGATTTCTTCTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGAAAGC 2058
Qy 2041 CGCATTCTCTGAGAACAGGTTAAAGAGTAAAGAGCTAGAGCCTGAGAGGAGAGAAACAGA 2100
Db 2059 CGCATTCTCTGAGAACAGGTTAAAGAGTAAAGAGCTAGAGCCTGAGAGGAGAGAAACAGA 2118
Qy 2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATATAATTCG 2160
Db 2119 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATATAATTCG 2178
Qy 2161 GAGCTCGAAGAGAAACATCGGAGGCGCAAGTCTCAGCCGAGCACCCTAGAGTGCACTG 2220
Db 2179 GAGCTCGAAGAGAAACATCGGAGGCGCAAGTCTCAGCCGAGCACCCTAGAGTGCACTG 2238
Qy 2221 AAACAGAAAGAGCAGCACTATAGGAGAAAGATTAAAGTGTGGACAAATCAGATATAAGAAA 2280
Db 2239 AAACAGAAAGAGCAGCACTATAGGAGAAAGATTAAAGTGTGGACAAATCAGATATAAGAAA 2298

Db 4459 AATAACAAACGAGCAGCAAGGCTGGGACAGGAAGTACATTGCTCTGGAGGATCAAAA 4518
QY GTCTCTATTATGACAAATGAGCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTGAG 4560
Db 4519 GTCTCTATTATGACAAATGAGCAGAGAGCTGGGACAGAGGCGGTGGAAGAAATTGAG 4578
QY CTGTGCTCTCCGACCGGGATGTATCTATTCAATGCTGCGTGTGCTTCCGAATCGCA 4620
Db 4579 CTGTGCTCTCCGACCGGGATGTATCTATTCAATGCTGCGTGTGCTTCCGAATCGCA 4638
QY AATACAGCCAAAGCAGAAAAGCAGAGCTGATGCTGCTTGAAGAACTCCCTGCTG 4680
Db 4639 AATACAGCCAAAGCAGAAAAGCAGAGCTGATGCTGCTTGAAGAACTCCCTGCTG 4698
QY AACTGGAAGGTGATGACCGCTTAGACATGAACTGACACGCTGCTTCAAGTACACAGGTG 4740
Db 4699 AACTGGAAGGTGATGACCGCTTAGACATGAACTGACACGCTGCTTCAAGTACACAGGTG 4758
QY GTGTGCTGGGACCGGAGGAGGCTTACGCGCTGAATGTCTTGAAGAACTCCCTAACC 4800
Db 4759 GTGTGCTGGGACCGGAGGAGGCTTACGCGCTGAATGTCTTGAAGAACTCCCTAACC 4818
QY CATGTCCAGGAATTTGGACAGCTTCTTCAAAATTTATATTATCAAGGACCTGGAAGCTA 4860
Db 4819 CATGTCCAGGAATTTGGACAGCTTCTTCAAAATTTATATTATCAAGGACCTGGAAGCTA 4878
QY CTCATGATAGCAGGAGAGCGGGCACTGTGTCTTGTGAGCGTGAAGAAAGTGAAACAG 4920
Db 4879 CTCATGATAGCAGGAGAGCGGGCACTGTGTCTTGTGAGCGTGAAGAAAGTGAAACAG 4938
QY TCCTGCGCCAGTCCCACTGCTGCGCAGCGGACATCTACCCAAATTTTGAAGCT 4980
Db 4939 TCCTGCGCCAGTCCCACTGCTGCGCAGCGGACATCTACCCAAATTTTGAAGCT 4998
QY GTCAAGGGGTGCGCACTTGTGTGGGAGGAGCAAGATTGAGAAAGGCTCTGATCTGTGCA 5040
Db 4999 GTCAAGGGGTGCGCACTTGTGTGGGAGGAGCAAGATTGAGAAAGGCTCTGATCTGTGCA 5058
QY GCATGCGCCAGCAAGTCTGATCTTCCGCTCAACGAAACCTCAGCAAAATCTGCAATC 5100
Db 5059 GCATGCGCCAGCAAGTCTGATCTTCCGCTCAACGAAACCTCAGCAAAATCTGCAATC 5118
QY CGGAAAGAGATAGAGACCTCAGAGCCCTGACGTGTATCCCTTACCAATTTACAGTATC 5160
Db 5119 CGGAAAGAGATAGAGACCTCAGAGCCCTGACGTGTATCCCTTACCAATTTACAGTATC 5178
QY CTCTTTGGAACCAATAAATTTCTACGAAATCGAATGAAGCAGTACACGCTCGAGGAATTC 5220
Db 5179 CTCTTTGGAACCAATAAATTTCTACGAAATCGAATGAAGCAGTACACGCTCGAGGAATTC 5238
QY CTGGATAAGAAATGACCAATTCCTTGGCACTGCTGTGTGTGCGCTCTTCCAAAGCTTC 5280
Db 5239 CTGGATAAGAAATGACCAATTCCTTGGCACTGCTGTGTGTGCGCTCTTCCAAAGCTTC 5298
QY CTGTCTCAATCTGAGTGAACAGCGCAGGAGCAGGAGAGTACTTGTGTGTTC 5340
Db 5299 CTGTCTCAATCTGAGTGAACAGCGCAGGAGCAGGAGAGTACTTGTGTGTTC 5358
QY CAGCAATTTGGAGTGTGTGATTTCTTACGGAAGCAGTACCGCAGAGAGTCTCAAG 5400
Db 5359 CAGCAATTTGGAGTGTGTGATTTCTTACGGAAGCAGTACCGCAGAGAGTCTCAAG 5418
QY TGGAGTGGCTTACCTTTGGCTTTTCCCTTACAGAGAACCTTCTTGTGTGACCCACTTC 5460
Db 5419 TGGAGTGGCTTACCTTTGGCTTTTCCCTTACAGAGAACCTTCTTGTGTGACCCACTTC 5478
QY AACTCAGTCAAGATTAATGAGATCAGGACGCTCTCAGCAGGAGAACCTTCCCGGAGCG 5520
Db 5479 AACTCAGTCAAGATTAATGAGATCAGGACGCTCTCAGCAGGAGAACCTTCCCGGAGCG 5538
QY TACCTGGACATCCCGAACCCCGGCTTACCTGGGCGCTGCAATTTCTCAGGAGGAGTATAC 5580
Db 5539 TACCTGGACATCCCGAACCCCGGCTTACCTGGGCGCTGCAATTTCTCAGGAGGAGTATAC 5598

QY 5581 TTGGGCTCTCATACCAAGGATAAATTAAAGGTCAATTTGCTGCAAGGAAACCTTCTGTAAG 5640
Db 5599 TTGGGCTCTCATACCAAGGATAAATTAAAGGTCAATTTGCTGCAAGGAAACCTTCTGTAAG 5658
QY 5641 GAGTCCGCACTGAACACCAACCGGGGCGGTCCACCTCCCGCAGCAGCCCAACCAAGCGA 5700
Db 5659 GAGTCCGCACTGAACACCAACCGGGGCGGTCCACCTCCCGCAGCAGCCCAACCAAGCGA 5718
QY 5701 GGCCCAACCAAGTAAACAGAGCACAATCAACAGCGCTGCGCTCCAGCCAGCGCCGCC 5760
Db 5719 GGCCCAACCAAGTAAACAGAGCACAATCAACAGCGCTGCGCTCCAGCCAGCGCCGCC 5778
QY 5761 GAAGGCCCAAGCCACCGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 5820
Db 5779 GAAGGCCCAAGCCACCGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 5838
QY 5821 GAGTGCAGGAGCAAGTCTCTGCGCGCCCTTGAGCGAGAGAGTCCCGCGCGCG 5880
Db 5839 GAGTGCAGGAGCAAGTCTCTGCGCGCCCTTGAGCGAGAGAGTCCCGCGCGCG 5898
QY 5881 ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTGAACAGCAGCAGCAGCGGC 5940
Db 5899 ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTGAACAGCAGCAGCAGCGGC 5958
QY 5941 CGGCTGCTCGGAGAGCGGTGAGGACCCCGCTGTCCAGGTGAACAGGAGAGGCGAG 6000
Db 5959 CGGCTGCTCGGAGAGCGGTGAGGACCCCGCTGTCCAGGTGAACAGGAGAGGCGAG 6018
QY 6001 AGTCCCTCTCAAGTTTTCAGGTTTAAACATGCTCACTTATGACTGGAATAAAGCTG 6060
Db 6019 AGTCCCTCTCAAGTTTTCAGGTTTAAACATGCTCACTTATGACTGGAATAAAGCTG 6078
QY 6061 GACAACTGCGCAGTAACTGCTCAGTCTGAGGATCATCCAGTGAATGGAGAAATCCGG 6120
Db 6079 GACAACTGCGCAGTAACTGCTCAGTCTGAGGATCATCCAGTGAATGGAGAAATCCGG 6138
QY 6121 CAGCAGTTGAAAGTCTGTTCTGAGAACAGATTATTC 6159
Db 6139 CAGCAGTTGAAAGTCTGTTCTGAGAACAGATTATTC 6177

RESULT 4
US-10-325-430-10
; Sequence 10, Application US/10325430
; Publication No. US2003015325A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE OF INVENTION: 32838,336 and 52908
; FILE REFERENCE: MPI01-294PIRNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-325-430-10

Query Match 100.0%; Score 6159; DB 15; Length 6574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTCAAGTTCAAATATGAGCGCGGATCCTTTGGATGCTGCTGCTGCAACCCATT 60
Db 19 ATGTTCAAGTTCAAATATGAGCGCGGATCCTTTGGATGCTGCTGCTGCAACCCATT 78

QY 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
DB |||||
QY 79 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACACCTTTATGACTCAA 138
DB |||||
QY 121 CAGCAGATGTCTCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
DB |||||
QY 139 CAGCAGATGTCTCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198
DB |||||
QY 181 GAATCAGTCAGCTCTCTGATGATGAATTAAGCAGCTGAGCACTTTGTCCGGAAGTAT 240
DB |||||
QY 199 GAATCAGTCAGCTCTCTGATGAATTAAGCAGCTGAGCACTTTGTCCGGAAGTAT 258
DB |||||
QY 241 TCCGACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAACTCAGA 300
DB |||||
QY 259 TCCGACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAACTCAGA 318
DB |||||
QY 301 AGCTTTGAGTTGTGTTCACTTTCTGTAAGTGAAGTGGTGAAGAGAAACACCGGG 360
DB |||||
QY 319 AGCTTTGAGTTGTGTTCACTTTCTGTAAGTGAAGTGGTGAAGAGAAACACCGGG 378
DB |||||
QY 361 GACATCTATGCTATGAAGTGAAGAGAGGCTTTATTGGCCCGAGGAGAGCTTTCA 420
DB |||||
QY 379 GACATCTATGCTATGAAGTGAAGAGAGGCTTTATTGGCCCGAGGAGAGCTTTCA 438
DB |||||
QY 421 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCACAAGCCCTGGATCCCCAATTA 480
DB |||||
QY 439 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCACAAGCCCTGGATCCCCAATTA 498
DB |||||
QY 481 CAGTATGCCCTTTCCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGG 540
DB |||||
QY 499 CAGTATGCCCTTTCCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGG 558
DB |||||
QY 541 GACTTGTCTCACTTTTGAATAGATATGAGGACCCAGTTAGATGAAAACCTGATCAGTTT 600
DB |||||
QY 559 GACTTGTCTCACTTTTGAATAGATATGAGGACCCAGTTAGATGAAAACCTGATCAGTTT 618
DB |||||
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAACGCTTATCTGATGGGATAGCTGATCGA 660
DB |||||
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTCAACGCTTATCTGATGGGATAGCTGATCGA 678
DB |||||
QY 661 GACATCAAGCCTTGAGAAACATTTCTGTTGACCCGACAGGACACATCAAGCTGGTGATTTT 720
DB |||||
QY 679 GACATCAAGCCTTGAGAAACATTTCTGTTGACCCGACAGGACACATCAAGCTGGTGATTTT 738
DB |||||
QY 721 GGATCTGCCCGGAAATGAATTAACAAGATGGTGAATGCCAACTCCCGATTGGGACC 780
DB |||||
QY 739 GGATCTGCCCGGAAATGAATTAACAAGATGGTGAATGCCAACTCCCGATTGGGACC 798
DB |||||
QY 781 CCAGATTATAGCTCTGAGTGCTGACTGTGATGAACGGGATGGAAGGACACCTTAC 840
DB |||||
QY 799 CCAGATTATAGCTCTGAGTGCTGACTGTGATGAACGGGATGGAAGGACACCTTAC 858
DB |||||
QY 841 GGCCTGGAATGTGACTGGTGGTCAGTGGGCGTGATTGCCCTATGAGATGATTTATGGGAGA 900
DB |||||
QY 859 GGCCTGGAATGTGACTGGTGGTCAGTGGGCGTGATTGCCCTATGAGATGATTTATGGGAGA 918
DB |||||
QY 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTCAGCGG 960
DB |||||
QY 919 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTCAGCGG 978
DB |||||
QY 961 TTTTTCGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1020
DB |||||
QY 979 TTTTTCGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1038
DB |||||
QY 1021 TTGTTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGTCATGCTTTCTTTC 1080
DB |||||
QY 1039 TTGTTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGTCATGCTTTCTTTC 1098
DB |||||
QY 1081 TCTAAAATTGACTGGACACATTCCTACTCTCTCCCGCTTCTGTTCCACCTCAAG 1140
DB |||||
QY 1099 TCTAAAATTGACTGGACACATTCCTACTCTCTCCCGCTTCTGTTCCACCTCAAG 1158
DB |||||

QY 1141 TCTGACGATGACACCTCCAATTTTGTATGAACCGAGAAAGAAATTCGTGGGTTTCATCTCT 1200
DB |||||
QY 1159 TCTGACGATGACACCTCCAATTTTGTATGAACCGAGAAAGAAATTCGTGGGTTTCATCTCT 1218
DB |||||
QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTCGCTTTTGTGGGTTTTCG 1260
DB |||||
QY 1219 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTCGCTTTTGTGGGTTTTCG 1278
DB |||||
QY 1261 TACAGCAAGCACTGGGGATTTCTTGTAGATCTGAGTCTGTGTTGTGCGGTCTGGACTCC 1320
DB |||||
QY 1279 TACAGCAAGCACTGGGGATTTCTTGTGTAGATCTGAGTCTGTGTTGTGCGGTCTGGACTCC 1338
DB |||||
QY 1321 CTTGCCAAGACTAGCTTCCATGGAAAGAAACTTCTCATCAAAAGCAAAAGCTTACAGAC 1380
DB |||||
QY 1339 CTTGCCAAGACTAGCTTCCATGGAAAGAAACTTCTCATCAAAAGCAAAAGCTTACAGAC 1398
DB |||||
QY 1381 TCTCAGACAAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGAGTGTCA 1440
DB |||||
QY 1399 TCTCAGACAAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGAGTGTCA 1458
DB |||||
QY 1441 GAGGTGAGGCTGTCTTACTAGTAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1500
DB |||||
QY 1459 GAGGTGAGGCTGTCTTACTAGTAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1518
DB |||||
QY 1501 TCCCTCTGAGCAGGACCTTGTCTACCTACATCACAGAATGCAGTAGCTTTAAAGCGAAGT 1560
DB |||||
QY 1519 TCCCTCTGAGCAGGACCTTGTCTACCTACATCACAGAATGCAGTAGCTTTAAAGCGAAGT 1578
DB |||||
QY 1561 TTGAGCAAGCAGCGGATGGAGGTGTCCAGGAGATGACAAAGCACTGCGAGCTTCTCCAT 1620
DB |||||
QY 1579 TTGAGCAAGCAGCGGATGGAGGTGTCCAGGAGATGACAAAGCACTGCGAGCTTCTCCAT 1638
DB |||||
QY 1621 GATATCAGAGCAGAGCCCGAGCTCCAGAAATCAAGAGCAGGAGTACCGAGCTCAA 1680
DB |||||
QY 1639 GATATCAGAGCAGAGCCCGAGCTCCAGAAATCAAGAGCAGGAGTACCGAGCTCAA 1698
DB |||||
QY 1681 GTGCAAGAAATGAGTTGATGAATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAAGA 1740
DB |||||
QY 1699 GTGCAAGAAATGAGTTGATGAATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAAGA 1758
DB |||||
QY 1741 CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGTTCAG 1800
DB |||||
QY 1759 CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGTTCAG 1818
DB |||||
QY 1801 CGGAAAGCGACAGAAATGTCAAGATAAATCTTGAAGCTAAGGATCAAGGGAAGCCTGAA 1860
DB |||||
QY 1819 CGGAAAGCGACAGAAATGTCAAGATAAATCTTGAAGCTAAGGATCAAGGGAAGCCTGAA 1878
DB |||||
QY 1861 GTGGGAGATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
DB |||||
QY 1879 GTGGGAGATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938
DB |||||
QY 1921 CTCGAGAGAACTGGAGAGGCTGCAAGGAGCGAGCCGAGAGGGAGCTGGAGAAGCTG 1980
DB |||||
QY 1939 CTCGAGAGAACTGGAGAGGCTGCAAGGAGCGAGCCGAGAGGGAGCTGGAGAAGCTG 1998
DB |||||
QY 1981 CAGAAACCGAGAGGATTTCTTCTGAAGGATCAGAAAGAAAGCTGGTGGAGCTGAGGAACGC 2040
DB |||||
QY 1999 CAGAAACCGAGAGGATTTCTTCTGAAGGATCAGAAAGAAAGCTGGTGGAGCTGAGGAACGC 2058
DB |||||
QY 2041 CGCATTTCTTGGAGAAACAAGTTAAAGACTAGAGACCATGAGACGCTAGAGAAACAGA 2100
DB |||||
QY 2059 CGCATTTCTTGGAGAAACAAGTTAAAGACTAGAGACCATGAGACGCTAGAGAAACAGA 2118
DB |||||
QY 2101 CTGAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2160
DB |||||
QY 2119 CTGAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2178
DB |||||
QY 2161 GAGCTCGAAGAGAAACATTCGGAGGCGCCAAAGTCTCAGCCAGCAGCCTAGAAAGTGCACCTG 2220
DB |||||
QY 2179 GAGCTCGAAGAGAAACATTCGGAGGCGCCAAAGTCTCAGCCAGCAGCCTAGAAAGTGCACCTG 2238
DB |||||
QY 2221 AAAACGAAAGAGACGACTATGAGGAAAGAGATTAAAGTGTGGCAANTCAGATTAAGAA 2280
DB |||||

Db	2239	AAACAGAAGAGCAGCACTATGAGAAAAGATTAAGTGTGGAACAATCAGATAAAGAAA	2298
Qy	2281	GACCTGGTGACAAGAGACACTGGAGAAATGATGCAGAGACACGAGGAGAGCGCCAT	2340
Db	2299	GACCTGGTGACAAGAGACACTGGAGAAATGATGCAGAGACACGAGGAGAGCGCCAT	2358
Qy	2341	GAGAAAGCCAAAATTTCTCAGCGAACAGAAAGCGATGATCAATGTATGGATTCCAGATC	2400
Db	2359	GAGAAAGCCAAAATTTCTCAGCGAACAGAAAGCGATGATCAATGTATGGATTCCAGATC	2418
Qy	2401	AGATCCCTGGAAACAGAGATTGTGGAACTGTCTGAAGCCATAAATTTGAGCAAAATAGC	2460
Db	2419	AGATCCCTGGAAACAGAGATTGTGGAACTGTCTGAAGCCATAAATTTGAGCAAAATAGC	2478
Qy	2461	AGTCTTTTATCCAAAGAAACATGAAGGCCCAAGAAGAGATGATTTCTGAATCAGGCAA	2520
Db	2479	AGTCTTTTATCCAAAGAAACATGAAGGCCCAAGAAGAGATGATTTCTGAATCAGGCAA	2538
Qy	2521	CAGAAATTTTACCTGGAGACACAGCTCGGGAAGTTGGAGGCCCAAGAACCTGAG	2580
Db	2539	CAGAAATTTTACCTGGAGACACAGCTCGGGAAGTTGGAGGCCCAAGAACCTGAG	2598
Qy	2581	GAGCAGCTGGAGAAGATCAGCCACCAAGACACACAGTGAACAAGATCGCTCTGGAATCG	2640
Db	2599	GAGCAGCTGGAGAAGATCAGCCACCAAGACACACAGTGAACAAGATCGCTCTGGAATCG	2658
Qy	2641	GAGACAGATTGCCGGAGGTCAGTCTAGAGCACGAGGACGAGAACTGGAGCTCAAGCGC	2700
Db	2659	GAGACAGATTGCCGGAGGTCAGTCTAGAGCACGAGGACGAGAACTGGAGCTCAAGCGC	2718
Qy	2701	CAGCTCACAGACTACAGCTCTCCCTCGAGGCGCAGTCAAGTTGACAGCCCTGCGAG	2760
Db	2719	CAGCTCACAGACTACAGCTCTCCCTCGAGGCGCAGTCAAGTTGACAGCCCTGCGAG	2778
Qy	2761	GCTGCACGGCGGCCCTGGAGAGCCAGCTTCCGCGGGAAGACAGAGCTGGGAAGAGACC	2820
Db	2779	GCTGCACGGCGGCCCTGGAGAGCCAGCTTCCGCGGGAAGACAGAGCTGGGAAGAGACC	2838
Qy	2821	ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAATTCAGCGC	2880
Db	2839	ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAATTCAGCGC	2898
Qy	2881	AAATTTGATGCTCTTCTTAACAGCTGTACTGTATACACAGCCTGGAGGACAGCTAAAC	2940
Db	2899	AAATTTGATGCTCTTCTTAACAGCTGTACTGTATACACAGCCTGGAGGACAGCTAAAC	2958
Qy	2941	CAGCTGACCGAGACACGCTGAACTCAACAAACAACTTCTACTGTTCACAGCACTC	3000
Db	2959	CAGCTGACCGAGACACGCTGAACTCAACAAACAACTTCTACTGTTCACAGCACTC	3018
Qy	3001	GATGAGGCTTCTGGCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTC	3060
Db	3019	GATGAGGCTTCTGGCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTC	3078
Qy	3061	CGCGGGAGATCAGGNAACGAGAGATGCGCTTACCAGCCAGAGCAACGATGGAGGCT	3120
Db	3079	CGCGGGAGATCAGGNAACGAGAGATGCGCTTACCAGCCAGAGCAACGATGGAGGCT	3138
Qy	3121	CTGAAGACACGCTGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAACGAT	3180
Db	3139	CTGAAGACACGCTGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAACGAT	3198
Qy	3181	GAGCTGCTGAAAAAGAGCGGCAAGTGGAGGCTCGAGAGCGCTCTGGGTGATGAGAAA	3240
Db	3199	GAGCTGCTGAAAAAGAGCGGCAAGTGGAGGCTCGAGAGCGCTCTGGGTGATGAGAAA	3258
Qy	3241	TCCAGATTTGAGTGTCCGGTTTCAGAGCTGACAGAAATCTCGACACCGAGAAACAGAGC	3300
Db	3259	TCCAGATTTGAGTGTCCGGTTTCAGAGCTGACAGAAATCTCGACACCGAGAAACAGAGC	3318
Qy	3301	AGGGCGAGAGCCGATCAGCGGATCACCGAGTCTGCCAGTCTGCGAGCTGGAGCTGGAGTGAAG	3360

Db	3319	AGGCGAGAGCGCGATCAAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG	3379
Qy	3361	GAGCAAAAGGCTGAGATCTCCGCTCTGCAGCAGGCTCTCAAAGAGCAGAAGCTGAAGGCC	3420
Db	3379	GAGCAAAAGGCTGAGATCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAAGCTGAAGGCC	3438
Qy	3421	GAGAGCCTCTCTGAACAAGCTCAATGA CTTGGAGAAGAAGCATGTA TGTCTTTGAAATGAAT	3480
Db	3439	GAGAGCCTCTCTGAACAAGCTCAATGA CTTGGAGAAGAAGCATGTA TGTCTTTGAAATGAAT	3498
Qy	3481	GCCCGAAGCTTTACAGCAGAAGCTGGAGACTGAACGAGAGCTCAAA CAGAGGCTTCTGGAA	3540
Db	3499	GCCCGAAGCTTTACAGCAGAAGCTGGAGACTGAACGAGAGCTCAAA CAGAGGCTTCTGGAA	3558
Qy	3541	GAGCAAGCCAAATTTACAGCAGCAGATGGACCTGCAGAAAAATCACAT TTTCCGCTCTGACT	3600
Db	3559	GAGCAAGCCAAATTTACAGCAGCAGATGGACCTGCAGAAAAATCACAT TTTCCGCTCTGACT	3618
Qy	3601	CAAGGACTGCAGAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAGAAGTGA CTTG	3660
Db	3619	CAAGGACTGCAGAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAGAAGTGA CTTG	3678
Qy	3661	GAGTATCAGCTGGAAAAACATTTCAAGTTCTCTATTCTCATGAAAAGTGA AAA TGGGAAGG	3720
Db	3679	GAGTATCAGCTGGAAAAACATTTCAAGTTCTCTATTCTCATGAAAAGTGA AAA TGGGAAGG	3738
Qy	3721	ACTATTCTTCAACAAACAACTCATATGATTTTCTGCAAGCCAAAAATGSA CCAACTGCT	3780
Db	3739	ACTATTCTTCAACAAACAAACCTCATATGATTTTCTGCAAGCCAAAAATGSA CCAACTGCT	3798
Qy	3781	AAAAAGAAAAAGGTTTATTATTGTCAGCGAAAAGAGACCTGCTTTTACCACACACAGGTT	3840
Db	3799	AAAAAGAAAAAGGTTTATTATTGTCAGCGAAAAGAGACCTGCTTTTACCACACACAGGTT	3858
Qy	3841	CCTCTGCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCTCGCTGTGCAGAG	3900
Db	3859	CCTCTGCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCTCGCTGTGCAGAG	3918
Qy	3901	CTAGAGAAAGCCCTTCAGAGAGCCGCATCGAGCTCCGGTCCGGCCCGCCGGAGGAAGCTGCC	3960
Db	3919	CTAGAGAAAGCCCTTCAGAGAGCCGCATCGAGCTCCGGTCCGGCCCGCCGGAGGAAGCTGCC	3978
Qy	3961	CACGCAAAACAA CGGACCA CCCCACACCCATCCAGCCAGCCACCGCAGGAGCAGAGATC	4020
Db	3979	CACGCAAAACAA CGGACCA CCCCACACCCATCCAGCCAGCCACCGCAGGAGCAGAGATC	4038
Qy	4021	GCCATGTCGCCATCGTGGGTTCGAGAGCACCGAGCCAGCCAGCTGCCATGAGCTGTGTGCC	4080
Db	4039	GCCATGTCGCCATCGTGGGTTCGAGAGCACCGAGCCAGCCAGCTGCCATGAGCTGTGTGCC	4098
Qy	4081	CCGCGCATCCAGCCGCGAAGAGAGTCTTCAACTCCAGAGGAATTTAGTCCGGGCTCTTAAG	4140
Db	4099	CCGCGCATCCAGCCGCGAAGAGAGTCTTCAACTCCAGAGGAATTTAGTCCGGGCTCTTAAG	4158
Qy	4141	GAACGATGACCAACAATATTTCTCAACGATTCACGATAGGACTGAA CANTGCGAGCCACA	4200
Db	4159	GAACGATGACCAACAATATTTCTCAACGATTCACGATAGGACTGAA CANTGCGAGCCACA	4218
Qy	4201	AAGTGTGCTGTGTCTGTGATACCGTGACTTTGGAGCGCCAGGCATCCAAATGTCTCGAA	4260
Db	4219	AAGTGTGCTGTGTCTGTGATACCGTGACTTTGGAGCGCCAGGCATCCAAATGTCTCGAA	4278
Qy	4261	TGTCAAGTATGATGTCA CCCCCAAGTGCTCCAGTGTTCAGCCAGCCACTGGGGCTTGCT	4320
Db	4279	TGTCAAGTATGATGTCA CCCCCAAGTGCTCCAGTGTTCAGCCAGCCACTGGGGCTTGCT	4338
Qy	4321	GCTGAATATGCCACACATTCACCGAGGCTTCTCCGCTGACAAATGCACTCCCCAGGT	4380
Db	4339	GCTGAATATGCCACACATTCACCGAGGCTTCTCCGCTGACAAATGCACTCCCCAGGT	4398
Qy	4381	CTCCAGCAAGAGAGCCCAAGCAGCAGCTTTGCA CTTGGAAGGTTGATGAAGGTGCCCAGG	4440
Db	4399	CTCCAGCAAGAGAGCCCAAGCAGCAGCTTTGCA CTTGGAAGGTTGATGAAGGTGCCCAGG	4458

4441 AATAAACAAGAGGACAGCAAGAGCTGGGACAGGAAGTACATTTGCTCGAGGGATCAAAA 4500
4459 AATAAACAAGAGGACAGCAAGAGCTGGGACAGGAAGTACATTTGCTCGAGGGATCAAAA 4518
4501 GTCCTCATTTATGACAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAAGATTTGAG 4560
4519 GTCCTCATTTATGACAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAAGATTTGAG 4578
4561 CTGTGCTTCCCGACGGGAGTATCTATTCATGCTGCGGTGGTCTTCGGAATCGCA 4620
4579 CTGTGCTTCCCGACGGGAGTATCTATTCATGCTGCGGTGGTCTTCGGAATCGCA 4638
4621 AATACAGCAAGAGCAAGAGAGCTGATGCTAACTGCTTGGAAATCTCCCTGCTG 4680
4639 AATACAGCAAGAGCAAGAGAGCTGATGCTAACTGCTTGGAAATCTCCCTGCTG 4698
4681 AAATCGAAGGTGATGACCGCTAGACATGAATCGACGCTGCCCTTCAGTGACAGGTG 4740
4699 AAATCGAAGGTGATGACCGCTAGACATGAATCGACGCTGCCCTTCAGTGACAGGTG 4758
4741 GTGTTGTTGGGACCGAGAGGCTCTACGCCCTGAATGCTTGAAGAACTCCCTAAC 4800
4759 GTGTTGTTGGGACCGAGAGGCTCTACGCCCTGAATGCTTGAAGAACTCCCTAAC 4818
4801 CATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTTATCAAGGACCTGGAGAAGCTA 4860
4819 CATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTTATCAAGGACCTGGAGAAGCTA 4878
4861 CTCATGATAGCAGAGAGAGCGGACATGTGTCTTTGTGGAGCTGAAGAAAGTGAACAG 4920
4879 CTCATGATAGCAGAGAGAGCGGACATGTGTCTTTGTGGAGCTGAAGAAAGTGAACAG 4938
4921 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTACCCCAATTTTGAAGCT 4980
4939 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTACCCCAATTTTGAAGCT 4998
4981 GTCAAGGGCTGCCACTTTTGGGAGCAAGATTTGAGAACGGCTCTGCACTGTGCA 5040
4999 GTCAAGGGCTGCCACTTTTGGGAGCAAGATTTGAGAACGGCTCTGCACTGTGCA 5058
5041 GCCATGCCAGCAAGTCTGATCTTCGCTACAGCAAGAACTCAGCAAAATCTGCACTC 5100
5059 GCCATGCCAGCAAGTCTGATCTTCGCTACAGCAAGAACTCAGCAAAATCTGCACTC 5118
5101 CGGAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTTACAGTATC 5160
5119 CGGAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTTACAGTATC 5178
5161 CTCATTGGAAACCAATAAATTTACGAAATCGACATGAAGCAGTACCGCTCAGGAATTC 5220
5179 CTCATTGGAAACCAATAAATTTACGAAATCGACATGAAGCAGTACCGCTCAGGAATTC 5238
5221 CTGGATAAGAAATCAACCATTTCTTGGCAGCTGCTGTTTGGCCCTCTTCCAAAGCTTC 5280
5239 CTGGATAAGAAATCAACCATTTCTTGGCAGCTGCTGTTTGGCCCTCTTCCAAAGCTTC 5298
5281 CCTGTCTCAATCTGCAGGTGAACACCGCAGGCGCAGCAGAGAGTACTGTGTGTTTC 5340
5299 CCTGTCTCAATCTGCAGGTGAACACCGCAGGCGCAGCAGAGAGTACTGTGTGTTTC 5358
5341 CACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGAGCTAGCGGACACCATCTCAAG 5400
5359 CACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGAGCTAGCGGACACCATCTCAAG 5418
5401 TGGAGTCTGCTTACCTTTGGCCCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC 5460
5419 TGGAGTCTGCTTACCTTTGGCCCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC 5478
5461 AACTCAGTGAAGTAAATTGAGATTCAGGCAAGCTCTCAGCAGGAGACCTTCCCGAGGG 5520
5479 AACTCAGTGAAGTAAATTGAGATTCAGGCAAGCTCTCAGCAGGAGACCTTCCCGAGGG 5538

5521 TACTTGACATCCGAAACCGCGCTACCTGGGCGCTGCGCATTTCTCAGAGCGATTTAC 5580
5539 TACTTGACATCCGAAACCGCGCTACCTGGGCGCTGCGCATTTCTCAGAGCGATTTAC 5598
5581 TTGGGCTCTCATACAGAGTAAATTAAGGTCATTTGCTGCAAGGAAACCTCTGTGAAG 5640
5599 TTGGGCTCTCATACAGAGTAAATTAAGGTCATTTGCTGCAAGGAAACCTCTGTGAAG 5658
5641 GAGTCCGGCACTCAACACCAACCGGGCCGCTCCACCTCCGCGAGCAGCCCAACAAAGCGA 5700
5659 GAGTCCGGCACTCAACACCAACCGGGCCGCTCCACCTCCGCGAGCAGCCCAACAAAGCGA 5718
5701 GGGCCACCCACGTCACAAACGAGCACAACCAAGCGGTGGCTTCAGCCTCAGCGCCGCC 5760
5719 GGGCCACCCACGTCACAAACGAGCACAACCAAGCGGTGGCTTCAGCCTCAGCGCCGCC 5778
5761 GAAGGCCACCGCCACCGCGAGAGCCACACCCACCGCTACCGCGAGGGCGGAC 5820
5779 GAAGGCCACCGCCACCGCGAGAGCCACACCCACCGCTACCGCGAGGGCGGAC 5838
5821 GAGTGGCGCAGGACAGTCTCTCTGGCCGCCCTCGAGCGAGAGAGTCCCGCGCGG 5880
5839 GAGTGGCGCAGGACAGTCTCTCTGGCCGCCCTCGAGCGAGAGAGTCCCGCGCGG 5898
5881 ATGCTCAGCAGCGGAGAGCGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGCGGC 5940
5899 ATGCTCAGCAGCGGAGAGCGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGCGGC 5958
5941 CGGCTGCTCGGGAGCGGTGAGACCCCGCTCTCCAGGTGAACAGGGAAGAGGGCAG 6000
5959 CGGCTGCTCGGGAGCGGTGAGACCCCGCTCTCCAGGTGAACAGGGAAGAGGGCAG 6018
6001 AGTGCCTCTCAAGTTTTCAGGTTTACGTTTACACTGTCACTTATGATGGAATAAAGCTG 6060
6019 AGTGCCTCTCAAGTTTTCAGGTTTACGTTTACACTGTCACTTATGATGGAATAAAGCTG 6078
6061 GACACCTGCGCAGTAACTGTCAGTCTCAGGATCATCCAGCTGAATGAGAAATCCGG 6120
6079 GACACCTGCGCAGTAACTGTCAGTCTCAGGATCATCCAGCTGAATGAGAAATCCGG 6138
6121 CAGCAGGTTGAAAGTCTGTTCTTGAGAACAGATTTATTC 6159
6139 CAGCAGGTTGAAAGTCTGTTCTTGAGAACAGATTTATTC 6177

RESULT 5
US-10-415-011-43
; Sequence 43, Application US/10415011
; Publication No. US20040053394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuming
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: NGUYEN, Darniel B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LAL, Preeti G.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: LU, Dying Alina M.
; APPLICANT: BOROWSKI, Mark L.


```
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/242,410
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/244,068
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/245,708
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 43
; LENGTH: 6298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7484496CB1
US-10-415-011-43

Query Match      92.0%; Score 5666.4; DB 13; Length 6298;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5961; Conservative 0; Mismatches 36; Indels 243; Gaps 3;

QY      1  ATCTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTGTCTGCTCAACCCATT 60
DB      55  ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTGTCTGCTCAACCCATT 114
QY      61  GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCACCTTTATGACTCAA 120
DB      115  GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCACCTTTATGACTCAA 174
QY      121  CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
DB      175  CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 234
QY      181  GAATGCACTGAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCCGGAAAGTAT 240
DB      235  GAATGCACTGAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCCGGAAAGTAT 294
QY      241  TCGCACACCATAGCTAGTTACAGAGCTCCAGCCTTCGGGAAAGGACTTGAAGTTCAGA 300
DB      295  TCGCACACCATAGCTAGTTACAGAGCTCCAGCCTTCGGGAAAGGACTTGAAGTTCAGA 354
QY      301  AGTCCTGTAGGTGTGGTCACTTTGCTGAAGTGCAGGTGGTGAAGAGAAAGCAACCGGG 360
DB      355  AGTCCTGTAGGTGTGGTCACTTTGCTGAAGTGCAGGTGGTGAAGAGAAAGCAACCGGG 414
QY      361  GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCCAAGAGAGGTTTCA 420
DB      415  GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCCAAGAGAGGTTTCA 474
QY      421  TTTTGTGAGGAGAGCGGAAACATATTATCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 480
DB      475  TTTTGTGAGGAGAGCGGAAACATATTATCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 534
QY      481  CAGTATGCCCTTTACGACAAAAAATCACCTTTATCTGATGGAGGAATATACGCTGGAGG 540

Db      535  CAGTATGCCCTTTACGACAAAAAATCACCTTTATCTGATGGAGGAATATACGCTGGAGG 594
QY      541  GACTTGTGTGTCACATTTTGAATAGATATGAGGACCAAGTTAGATGAAAAACCTGATACAGTTT 600
DB      595  GACTTGTGTGTCACATTTTGAATAGATATGAGGACCAAGTTAGATGAAAAACCTGATACAGTTT 654
QY      601  TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCTTCAATCTGATGGGATAGCTGCATCGA 660
DB      655  TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCTTCAATCTGATGGGATAGCTGCATCGA 714
QY      661  GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGCTGGTGGATTT 720
DB      715  GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGCTGGTGGATTT 774
QY      721  GGATCTGCGCGGAAAAATGAATTCAAAAAAGATGTGAATGCCAAACTCCCGATTTGGGACC 780
DB      775  GGATCTGCGCGGAAAAATGAATTCAAAAAAGATGTGAATGCCAAACTCCCGATTTGGGACC 834
QY      781  CCAGATTACATGGCTCCTGAGTGTCTGATGTAACGGGATGAAAAAGGCACTAC 840
DB      835  CCAGATTACATGGCTCCTGAGTGTCTGATGTAACGGGATGAAAAAGGCACTAC 894
QY      841  GGCCTGGAATGTGACTGGTGTGCTGAGTGGGCTGATTTGCCCTATGAGATGATTTATGGAGA 900
DB      895  GGCCTGGAATGTGACTGGTGTGCTGAGTGGGCTGATTTGCCCTATGAGATGATTTATGGAGA 954
QY      901  TCCCCCTTCCGAGAGGAACTCTGCCAGAACCTTCAATAACATTAATGATTTCCAGCG 960
DB      955  TCCCCCTTCCGAGAGGAACTCTGCCAGAACCTTCAATAACATTAATGATTTCCAGCG 1014
QY      961  TTTTGTGAAATTTCCAGATGACCCCAAGTGTGATGACTTTCTTGATCTGATTTCAAGC 1020
DB      1015  TTTTGTGAAATTTCCAGATGACCCCAAGTGTGATGACTTTCTTGATCTGATTTCAAGC 1074
QY      1021  TTGTTGTGCGCCAGAAAGAGAGACTGAAAGTTTGAAGGTCTTTGCTGCGCACTCTTCTTC 1080
DB      1075  TTGTTGTGCGCCAGAAAGAGAGACTGAAAGTTTGAAGGTCTTTGCTGCGCACTCTTCTTC 1134
QY      1081  TCTAAAAATGACTGGAACAACATTCGTAACCTCTCCCTCCCTCTGTTCCACCCCTCAAG 1140
DB      1135  TCTAAAAATGACTGGAACAACATTCGTAACCTCTCTCCCTCTGTTCCACCCCTCAAG 1194
QY      1141  TCTGACGATGACACCTCCAAATTTTGATGAACCAAGAGAAATTCGTGGGTTTCATCTCT 1200
DB      1195  TCTGACGATGACACCTCCAAATTTTGATGAACCAAGAGAAATTCGTGGGTTTCATCTCT 1254
QY      1201  CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTGCGGTTGTTGGGGTTTCG 1260
DB      1255  CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTGCGGTTGTTGGGGTTTCG 1314
QY      1261  TACAGCAAGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTTGGGCTCTGGACTCC 1320
DB      1315  TACAGCAAGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTTGGGCTCTGGACTCC 1374
QY      1321  CTTGCCAAGACTAGCTCCATGGAAGAAATCTCTCATCAAAAGCAAGAGCTTACAGAC 1380
DB      1375  CTTGCCAAGACTAGCTCCATGGAAGAAATCTCTCATCAAAAGCAAGAGCTTACAGAC 1434
QY      1381  TCTCAGCAAGTGTCACAAGATGGAGAGAAATGACCCGGTTACATCGGAGAGTGTCA 1440
DB      1435  TCTCAGCAAGTGTCACAAGATGGAGAGAAATGACCCGGTTACATCGGAGAGTGTCA 1494
QY      1441  GAGTGTGAGGCTGTGCTTAGTCTAGAGAGGAGTGGAGCTGAGGCTCTGAGACTTACAGA 1500
DB      1495  GAGTGTGAGGCTGTGCTTAGTCTAGAGAGGAGTGGAGCTGAGGCTCTGAGACTTACAGA 1554
QY      1501  TCCTCTCTGAGCAGGACCTTGTCTACTCATCACAGATGCAAGTCTTAAAGCAAGT 1560
DB      1555  TCCTCTCTGAGCAGGACCTTGTCTACTCATCACAGATGCAAGTCTTAAAGCAAGT 1614
QY      1561  TTGGAGCAAGCAAGGATGGAGGTGTTCCAGGAGGATGCAAAAGCACTGAGCTTCTCCAT 1620
```


Db 1615 TTGGAGCAAGCACCGATGGAGGTGTCCACGAGGATGACAAAGCACATCGAGCTTCTCCAT 1674
Qy 1621 GATATCAGAGACGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACACAGCTCAA 1680
Db 1675 GATATCAGAGACGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACAGGCTCAA 1734
Qy 1681 GTGGAAGAAATCAGGTTGATGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740
Db 1735 GTGGAAGAAATCAGGTTGATGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAGA 1794
Qy 1741 CGGAGTGATCTCTAGAAATCTGAGCTGAGAGAGTCTCGGTTGCTGTGAAGATTCAG 1800
Db 1795 CGGAGTGATCTCTAGAAATCTGAGCTGAGAGAGTCTCGGTTGCTGTGAAGATTCAG 1854
Qy 1801 CGGAAAGCGACAGAATGTCAAGATAAATCTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1860
Db 1855 CGGAAAGCGACAGAATGTCAAGATAAATCTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1914
Qy 1861 GTGGAGAAATATGCGAAATCTGAGAAAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAG 1920
Db 1915 GTGGAGAAATATGCGAAATCTGAGAAAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAG 1974
Qy 1921 CTCGAAGGAACTCGAAGAGCTG----- 1945
Db 1975 CTCGAAGGAACTCGAAGAGCTGTAAAGCCAGCACGGAGGCCACCGAGCTGCTGCAG 2034
Qy 1946 -----CAAAGAGAGCCGAGAGGGAGCTGGAGAGCTGGAGAGCTCGAAGCTCGAAGCCGAGAG 1992
Db 2035 AATATCCGCGCAAGGAGAGCGAGCGAGAGGGAGCTGGAGAGCTGCAGAAACCGAGAG 2094
Qy 1993 GATTCTTCTGAAGGATCAGAAAGAGCTGGTGGAGAGCTGAGGAACCGCGCATCTCTG 2052
Db 2095 GATTCTTCTGAAGGATCAGAAAGAGCTGGTGGAGAGCTGAGGAACCGCGCATCTCTG 2154
Qy 2053 GAGAAACAAGGTAAAGAGACTAGAGACCACTGGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2112
Db 2155 GAGAACAGGTAAAGAGACTAGAGACCACTGGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2214
Qy 2113 ATCCAGACAAAATCCAAACAGATCCAGCAGATGGCTGATATAAATCTCGAGCTCGAAGAG 2172
Db 2215 ATCCAGACAAAATCCAAACAGATCCAGCAGATGGCTGATATAAATCTCGAGCTCGAAGAG 2274
Qy 2173 AAACATCGGGAGGCCCAAGCTCTCAGCCAGCAGCTAGAGAGTCACTGAAACAGAAAGAG 2232
Db 2275 AAACATCGGGAGGCCCAAGCTCTCAGCCAGCAGCTAGAGAGTCACTGAAACAGAAAGAG 2334
Qy 2233 CAGCACTATGAGGAAAGATTAAAGTGTGGACAATCAGATATAAGAAAGACCTGCTGCAC 2292
Db 2335 CAGCACTATGAGGAAAGATTAAAGTGTGGACAATCAGATATAAGAAAGACCTGCTGCAC 2394
Qy 2293 AAGGAGACACTGGAGAACATGATGACAGACACGAGAGAGGCCCATGAGAGGCGCAA 2352
Db 2395 AAGGAGACACTGGAGAACATGATGACAGACACGAGAGAGGCCCATGAGAGGCGCAA 2454
Qy 2353 ATTCTCAGCGAAACAGAAAGCGATGATCAATGTATGGATTCCAAGATCAGATCCCTGGAA 2412
Db 2455 ATTCTCAGCGAACAGAAAGCGATGATCAATGTATGGATTCCAAGATCAGATCCCTGGAA 2514
Qy 2413 CAGAGATTGTGAACTGTCTGAGCCCAATAAATCTTGACGAAATAGCAGTCTTTTACC 2472
Db 2515 CAGAGATTGTGAACTGTCTGAGCCCAATAAATCTTGACGAAATAGCAGTCTTTTACC 2574
Qy 2473 CAAAGGAACATGAAAGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2532
Db 2575 CAAAGGAACATGAAAGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2634
Qy 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAAACTGGAGAGCAGCTGGAG 2592
Db 2635 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAAACTGGAGAGCAGCTGGAG 2694
Qy 2593 AAGATCAGCCACCAAGACCACTGACAGAAATCGGCTGCTGGAACCTGGAGACAGATTG 2652
Db 2695 AAGATCAGCCACCAAGACCACTGACAGAAATCGGCTGCTGGAACCTGGAGACAGATTG 2754

Qy 2653 CGGAGGTCTAGAGACGAGGAGCGAAACTGGAGCTCAAGCGCCAGCTCAGAG 2712
Db 2755 CGGAGGTCTAGAGACGAGGAGCGAAACTGGAGCTCAAGCGCCAGCTCAGAG 2814
Qy 2713 CTAAGCTTCTCCCTGAGAGCGCGAGTCAAGTTGACAGCCCTGACGGCTGACGGCG 2772
Db 2815 CTAAGCTTCTCCCTGAGGAGCGCGAGTCAAGTTGACAGCCCTGACGGCTGACGGCG 2874
Qy 2773 GCCTGAGAGCGCAGCTTCCGACGAGGCGAAGACAGAGCTGGAAGAGACCAAGAGCT 2832
Db 2875 GCCTGAGAGCGCAGCTTCCGACGAGGCGAAGACAGAGCTGGAAGAGACCAAGAGCT 2934
Qy 2833 GAAAGAGAGATCAAGGCACTCAAGGCAATAGAGATGAATCCAGCGCAAAATTTGATGCT 2892
Db 2935 GAAAGAGAGATCAAGGCACTCAAGGCAATAGAGATGAATCCAGCGCAAAATTTGATGCT 2994
Qy 2893 CTTCTGAAGCTGTACTGTAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 2952
Db 2995 CTTCTGAAGCTGTACTGTAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 3054
Qy 2953 GACAACTGAACTCAACAAACCAAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3012
Db 3055 GACAACTGAACTCAACAAACCAAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3114
Qy 3013 GCGCCAAACAGAGAGTTGTACAACTCGGAAGTGAAGTGAAGCTCTCCGCGGGAGATC 3072
Db 3115 GCGCCAAACAGAGAGTTGTACAACTCGGAAGTGAAGTGAAGCTCTCCGCGGGAGATC 3174
Qy 3073 ACGGAAACAGAGATGCAAGCTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3132
Db 3175 ACGGAAACAGAGATGCAAGCTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3234
Qy 3133 TGCACCAATGTGAGAGAAACAGGCTCATGGATTTGGAGCCCTTAAACGATGAGCTGTAGAA 3192
Db 3235 TGCACCAATGTGAGAGAAACAGGCTCATGGATTTGGAGCCCTTAAACGATGAGCTGTAGAA 3294
Qy 3193 AAGAGCGGAGTGGAGGCTCGAGAGAGCTCTCGGCTGATGAGAATCCAGTTTGGAG 3252
Db 3295 AAGAGCGGAGTGGAGGCTCGAGAGAGCTCTCGGCTGATGAGAATCCAGTTTGGAG 3354
Qy 3253 TGTCCGTTTCCAGAGCTGACAGAAATGCTGGACAACGAGAAACAGAGCGGCGAGAGCC 3312
Db 3355 TGTCCGTTTCCAGAGCTGACAGAAATGCTGGACAACGAGAAACAGAGCGGCGAGAGCC 3414
Qy 3313 GATCAGCGGATCACCGAGTCTCCAGAGTGGAGCTGGAGCTGAGGAGGACCAAGGCT 3372
Db 3415 GATCAGCGGATCACCGAGTCTCCAGAGTGGAGCTGGAGCTGAGGAGGACCAAGGCT 3474
Qy 3373 GAGATTCTCGCTCTGACAGAGGCTCTCAAGAGCAGAAAGCTGAAAGCGGAGAGCTCTCT 3432
Db 3475 GAGATTCTCGCTCTGACAGAGGCTCTCAAGAGCAGAAAGCTGAAAGCGGAGAGCTCTCT 3534
Qy 3433 GACAACTCAATGACCTGGAGAAAGCATGCTTATGCTTGAATGAATGCTCCGAGCTTA 3492
Db 3535 GACAACTCAATGACCTGGAGAAAGCATGCTTATGCTTGAATGAATGCTCCGAGCTTA 3594
Qy 3493 CAGCAGAGCTGGAGCTGAAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGCCAAA 3552
Db 3595 CAGCAGAGCTGGAGCTGAAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGCCAAA 3654
Qy 3553 TTACAGCAGCAGATGGAGCTTCAAGAAATCACAATTTTCCGCTGCTGACTCAAGGCTGAA 3612
Db 3655 TTACAGCAGCAGATGGAGCTTCAAGAAATCACAATTTTCCGCTGCTGACTCAAGGCTGAA 3714
Qy 3613 GAACTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGGAGTATCAGCTG 3672
Db 3715 GAACTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGGAGTATCAGCTG 3774
Qy 3673 GAAACCAATCAGGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3732
Db 3775 GAAACCAATCAGGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3834

Qy 3733 CAAACCAAACTCATTTGTTCTTCTGCAAGCCAAATGAGACCAACTCTGCTAAAAAGAAAAAG 3792
Db |||||
Qy 3835 CAAACCAAACTCATTTGTTCTTCTGCAAGCCAAATGAGACCAACTCTGCTAAAAAGAAAA-- 3892
Db |||||
Qy 3793 GGTATTATTAGTCGACGGAAAGAGGACCTGCTTTTACCACACAGGTTCTCTGCGAGTAC 3852
Db |||||-----AGTTCTCTGCGAGTAC 3909
Qy 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTTAGAGGAAGCC 3912
Db |||||
Qy 3910 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTTAGAGGAAGCC 3969
Qy 3913 CTTTCAGAAAGACCCGATCGAGCTCCGCTCGCCCGGGAGGAAGCTGCCACCGCAAGACA 3972
Db |||||
Qy 3970 CTTTCAGAAAGACCCGATCGAGCTCCGCTCGCCCGGGAGGAAGCTGCCACCGCAAGACA 4029
Qy 3973 ACGACCAACCCACACCCATCCACGCCAGCCACCGAGGAGGAGAGATCGCCATGTCGCCG 4032
Db |||||
Qy 4030 ACGACCAACCCACACCCATCCAGCCAGCCACCGAGGAGGAGATCGCCATGTCGCCG 4089
Qy 4033 ATCGTCGGTGGCCAGAGCACCAAGCCAGTGCATGAGCCTGTGTCGCCCGCCATCCAGC 4092
Db |||||
Qy 4090 ATCGTCGGTGGCCAGAGCACCAAGCCAGTGCATGAGCCTGTGTCGCCCGCCATCCAGC 4149
Qy 4093 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCCGCGTCTTAAGGAAAGCATGAC 4152
Db |||||
Qy 4150 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCCGCGTCTTAAGGAAAGCATGAC 4209
Qy 4153 CACAATATTTCTTCAACGATTTCAACGTTAGGACTGAACAATCGAGGCCACAAAGTGTCTGTG 4212
Db |||||
Qy 4210 CACAATATTTCTTCAACGATTTCAACGTTAGGACTGAACAATCGAGGCCACAAAGTGTCTGTG 4269
Qy 4213 TGTCTGGATACCGTGACATTTGAGCCGAGGATCCAAATGCTCGAATGTGAGGTGATG 4272
Db |||||
Qy 4270 TGTCTGGATACCGTGACATTTGAGCCGAGGATCCAAATGCTCGAATGTGAGGTGATG 4329
Qy 4273 TGTCAACCCCAAGTGTCTTCCAGCTTGTGCCAGCCACTCGGGTGTGCTCTGCAATATGCC 4332
Db |||||
Qy 4330 TGTCAACCCCAAGTGTCTTCCAGCTTGTGCCAGCCACTCGGGTGTGCTCTGCAATATGCC 4389
Qy 4333 ACACATTTCAACGAGGCTTGTCCGTGACAAATGAATCCGCCAGGTTCAGACCAAG 4392
Db |||||
Qy 4390 ACACATTTCAACGAGGCTTGTCCGTGACAAATGAATCCGCCAGGTTCAGACCAAG 4449
Qy 4393 GAGCCAGAGAGCTTGACCTGGAGGAGGTGATGAGGTGCCAGGATTAACAAACGA 4452
Db |||||
Qy 4450 GAGCCAGAGAGCTTGACCTGGAGGAGGTGATGAGGTGCCAGGATTAACAAACGA 4509
Qy 4453 GACAGCAAGGCTGGAGAGGATGATGTTCTGGAGGAGTCAAAAGTCTCTATTAT 4512
Db |||||
Qy 4510 GACAGCAAGGCTGGAGAGGATGATGTTCTGGAGGAGTCAAAAGTCTCTATTAT 4569
Qy 4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCCTCCC 4572
Db |||||
Qy 4570 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCCTCCC 4629
Qy 4573 GAGGGGATGATATCTATTCATGCTGCGCTTGTGCTTCCGAACTCGCAAAATACAGCCAAA 4632
Db |||||
Qy 4630 GAGGGGATGATATCTATTCATGCTGCGCTTGTGCTTCCGAACTCGCAAAATACAGCCAAA 4689
Qy 4633 GCA----- 4635
Db |||||
Qy 4690 GCAGATGTCCCATACATACCTGAAGATGGAATCTCACCGCACACACCTGCTGGCCCGG 4749
Qy 4636 ----- 4635
Db |||||
Qy 4750 AGAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4809
Qy 4636 ----- 4635
Db |||||
Qy 4810 GAATCAGTTGTCCAGGTGGAGAGTTTCTAGGAAAAAGAGAGCTGATCTAACTG 4662
Db |||||
Qy 4663 CTTGGAAACTCTCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTG 4722

Db 4870 CTTGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTG 4929
Qy 4723 CCCTTCAGTGAAGCAGGTGGTGTGGTGGGCACCGAGGAGGGCTCTAGCGCCTGAATGTC 4782
Db 4930 CCCCTTCAGTGAAGCAGGTGGTGTGGTGGGCACCGAGGAGGGCTCTAGCGCCTGAATGTC 4989
Qy 4783 TTGAAAACTCCCTTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATTC 4842
Db 4990 TTGAAAACTCCCTTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATTC 5049
Qy 4843 AAGAACCTTGAAGAGCTTACTATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 4902
Db 5050 AAGAACCTTGAAGAGCTTACTATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 5109
Qy 4903 GTGAAGAAAGTGAACCAAGTCCCTGGCCCAAGTCCCACTGCTGCCAGCCCGACATCTCA 4962
Db 5110 GTGAAGAAAGTGAACCAAGTCCCTGGCCCAAGTCCCACTGCTGCCAGCCCGACATCTCA 5169
Qy 4963 CCCAACATTTTGAAGCTGTCAAGGGTGCACCTTGTTCGGGAGGCAAGATTTGAAC 5022
Db 5170 CCCAACATTTTGAAGCTGTCAAGGGTGCACCTTGTTCGGGAGGCAAGATTTGAAC 5229
Qy 5023 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTGCAATTTCTCGCTACACGAAAC 5082
Db 5230 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTGCAATTTCTCGCTACACGAAAC 5289
Qy 5083 CTCAGCAAAATCTGCATCTCCGAAAGAGATAGACCTTCAGAGCCCTGCAGCTGTATCCAC 5142
Db 5290 CTCAGCAAAATCTGCATCTCCGAAAGAGATAGACCTTCAGAGCCCTGCAGCTGTATCCAC 5349
Qy 5143 TTCAACCAATTTACAGTATCTCTTGGAAAGTGAACCAATTTCTCGAAATTCGACATGAAGCAG 5202
Db 5350 TTCAACCAATTTACAGTATCTCTTGGAAAGTGAACCAATTTCTCGAAATTCGACATGAAGCAG 5409
Qy 5203 TACACGCTCCAGGAATTTCTGGATAGAAATGAACCAATTTCTCGCTGGCACTGTGTGTTGCC 5262
Db 5410 TACACGCTCCAGGAATTTCTGGATAGAAATGAACCAATTTCTCGCTGGCACTGTGTGTTGCC 5469
Qy 5263 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGAGGCGAGAG 5322
Db 5470 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGAGGCGAGAG 5529
Qy 5323 GAGTACTTGTGTGTTTCCAGCAATTTGGAGTGTGTGTGATTTCTTACGGAAGACGTAGC 5382
Db 5530 GAGTACTTGTGTGTTTCCAGCAATTTGGAGTGTGTGTGATTTCTTACGGAAGACGTAGC 5589
Qy 5383 CGCACAGACGATCTCAAGTGGAGTGTCTTACCTTTGGCCCTTTGCCCTTACAGAGAACCTTAT 5442
Db 5590 CGCACAGACGATCTCAAGTGGAGTGTCTTACCTTTGGCCCTTTGCCCTTACAGAGAACCTTAT 5649
Qy 5443 CTGTTTGTGACCCACTTCAACTCAGTCCGAGTAAATTTGAGATCCAGGACGCTCCTCAGCA 5502
Db 5650 CTGTTTGTGACCCACTTCAACTCAGTCCGAGTAAATTTGAGATCCAGGACGCTCCTCAGCA 5709
Qy 5503 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAAACCGCGCTTACCTGGGCGCTTGCATT 5562
Db 5710 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAAACCGCGCTTACCTGGGCGCTTGCATT 5769
Qy 5563 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGAGTAAATTAAGGTCATTTGCTGC 5622
Db 5770 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC 5829
Qy 5623 AAGGAAACCTCTGTGAAGGAGTCCGGCACTGAACACACCGGGGCGCTTCCACCTCCCGC 5682
Db 5830 AAGGAAACCTCTGTGAAGGAGTCCGGCACTGAACACACCGGGGCGCTTCCACCTCCCGC 5889
Qy 5683 AGCAGCCCAAGAGGAGGCGCCACCAAGTACAGAGACATACCAAGGCGCTGGCC 5742
Db 5890 AGCAGCCCAAGAGGAGGCGCCACCAAGTACAGAGACATACCAAGGCGCTGGCC 5949
Qy 5743 TCAGAGCCAGCGCGCCGAGGCGCCAGCCAGCGAGGCGCAAGCACCCACCGC 5802

Db 5950 TCCAGCCAGCGCCGCCGAGGCCCCAGCCACCAGCCGAGAGCCCAAGCACACCCACCGC 6009
Qy 5803 TACCCGAGGGCGGACCGAGCTGCGCAGGACAGTCTCTGCGCGCCCTCGAGGGA 5862
Db 6010 TACCCGAGGGCGGACCGAGCTGCGCAGGACAGTCTCTGCGCGCCCTCGAGGGA 6069
Qy 5863 GAGAAAGTCCCCGGCGGATGCTCAGCACGCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 5922
Db 6070 GAGAAAGTCCCCGGCGGATGCTCAGCACGCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 6129
Qy 5923 GAAGCAGCAGCAGGGCGCGGTGCGCTGCGGAGCGGTGAGGACCCCGCTGTCAGGTG 5982
Db 6130 GAAGCAGCAGCAGGGCGCGGTGCGCTGCGGAGCGGTGAGGACCCCGCTGTCAGGTG 6189
Qy 5983 AACCAAGGAGAGGCGAGAGTCCCTCTCAAGTTTTCAGGTTTAAACATGTCACCTATTAT 6042
Db 6190 AACCAAGGCTGGACCAAGTCTTCAGTATAAATCTCAGCCAGAGAAAACCAATCTCACT 6249

RESULT 6
US-10-028-946-1
; Sequence 1, Application US/10028946
; Publication No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6165
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-028-946-1

Query Match 91.9%; Score 5661.4; DB 14; Length 6165;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

Qy 1 ATGTTGAAGTTCAAATATGAGCGGGGAATCCTTTGGATGCTGGTGTCTGCTGAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGGGGAATCCTTTGGATGCTGGTGTCTGCTGAACCCATT 60
Qy 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACCCACCTTTATGACTCAA 120
Db 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACCCACCTTTATGACTCAA 120
Qy 121 CAGCAGATGTCCTCTCTTTCCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTTTTGA 180
Db 121 CAGCAGATGTCCTCTCTTTCCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTTTTGA 180
Qy 181 GAATGAGTCAAGCTCTCTGATGAAGATTAGCAGCTGAGCAACTTTGTCGGGAAGTAT 240
Db 181 GAATGAGTCAAGCTCTCTGATGAAGATTAGCAGCTGAGCAACTTTGTCGGGAAGTAT 240
Qy 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 300
Qy 301 AGTCCTTGTAGTTGTGGTCACTTTGCTGAAGTGCAGGTGTAAGAGAGAAAGCAACCGGG 360
Db 301 AGTCCTTGTAGTTGTGGTCACTTTGCTGAAGTGCAGGTGTAAGAGAGAAAGCAACCGGG 360
Qy 361 GACATCTTGTCTAAGTATGAAGAGAGAGGCTTTTATGGCCAGGAGCAGGTTTCA 420
Db 361 GACATCTTGTCTAAGTATGAAGAGAGAGGCTTTTATGGCCAGGAGCAGGTTTCA 420

Qy 421 TTTTGTGAGGAGAGGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGGAGAGGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480
Qy 481 CAGTATGCCCTTTCAGGACAAAATTCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
Db 481 CAGTATGCCCTTTCAGGACAAAATTCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
Qy 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGTATACAGTTT 600
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGTATACAGTTT 600
Qy 601 TACCTAGCTGAGTGTATTTTGGTGTTCACAGCGTTTCACTGATGGGATACGTGTCATCGA 660
Db 601 TACCTAGCTGAGTGTATTTTGGTGTTCACAGCGTTTCACTGATGGGATACGTGTCATCGA 660
Qy 661 GACATCAAGCCTCAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
Db 661 GACATCAAGCCTCAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
Qy 721 GGATCTCGCGAAATGAATTCAAACAAGATGGTGAATCCCAACTCCCGATTTGGAGC 780
Db 721 GGATCTCGCGAAATGAATTCAAACAAGATGGTGAATCCCAACTCCCGATTTGGAGC 780
Qy 781 CCAGATTACATGCTCCTGAAAGTCTGACTGTGATGAACGGGATGGAAGACACCTAC 840
Db 781 CCAGATTACATGCTCCTGAAAGTCTGACTGTGATGAACGGGATGGAAGACACCTAC 840
Qy 841 GGCCTGACTGTGACTGGTGTGATGCTGGGCTGATTCCTTATGAGATGATTTATGGAGA 900
Db 841 GGCCTGACTGTGACTGGTGTGATGCTGGGCTGATTCCTTATGAGATGATTTATGGAGA 900
Qy 901 TCCCTTTCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATGAAATTCACGGG 960
Db 901 TCCCTTTCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATGAAATTCACGGG 960
Qy 961 TTTTGTGAATTTCCAGATGATCCCAAGTGCAGAGTCACTTTCTGATCTGATTCAAAGC 1020
Db 961 TTTTGTGAATTTCCAGATGATCCCAAGTGCAGAGTCACTTTCTGATCTGATTCAAAGC 1020
Qy 1021 TTGTTGTGCGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
Db 1021 TTGTTGTGCGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
Qy 1081 TCTAAATTTGACTGGAACAACTTCGTAACCTCTCCCTCCCTTCCTCCACCCCTCAAG 1140
Db 1081 TCTAAATTTGACTGGAACAACTTCGTAACCTCTCTCCCTTCCTCCACCCCTCAAG 1140
Qy 1141 TCTCAGCATGACACCTCCAAATTTTGTATGAACCCAGAGAGAAATTCGTGGGTTTCATCTCT 1200
Db 1141 TCTCAGCATGACACCTCCAAATTTTGTATGAACCCAGAGAGAAATTCGTGGGTTTCATCTCT 1200
Qy 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTGCCGTTTGTGGGTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTGCCGTTTGTGGGTTTTCG 1260
Qy 1261 TACAGCAAGGCACTGGGATTTCTTGGTAGATCTGAGTCTGTTGTCGGGTCTGACTCC 1320
Db 1261 TACAGCAAGGCACTGGGATTTCTTGGTAGATCTGAGTCTGTTGTCGGGTCTGACTCC 1320
Qy 1321 CCTGCCAAGACTAGCTCCATGGAAAAAATACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
Db 1321 CCTGCCAAGACTAGCTCCATGGAAAAAATACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
Qy 1381 TCTCAGCAAGTGTCAAGAGTGGACGAAATGACCCGTTTACATCGGAGGTCTCA 1440
Db 1381 TCTCAGCAAGTGTCAAGAGTGGACGAAATGACCCGTTTACATCGGAGGTCTCA 1440
Qy 1441 GAGTGGAGGCTGTCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1441 GAGTGGAGGCTGTCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Qy 1501 TCCCTCTCGAGGAGGAGCTTGTCTACCTACATCACAGANTGCAGTACGTAAAGCGAGT 1560

1501 TCCTCTCGAGCAGGACCTTGTACCTATACACAGAAATGCAAGTAAAGCGAAGT 1560
1561 TTGGAGCAAGCA CGGATGAGGTGTCCAGCAGGATGACAAAGCACTGCGAGTCTTCCAT 1620
1561 TTGGAGCAAGCA CGGATGAGGTGTCCAGCAGGATGACAAAGCACTGCGAGTCTTCCAT 1620
1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGGAGTACCAAGGCTCAA 1680
1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGGAGTACCAAGGCTCAA 1680
1681 GTGGAAGAAATGAGGTGATGATGAATCAAGTTCGGAAGAGATCTTCTCAGCAGAGGA 1740
1681 GTGGAAGAAATGAGGTGATGATGAATCAAGTTCGGAAGAGATCTTCTCAGCAGAGGA 1740
1741 CGGAGTGAATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGCTGCTGAAGAAATTCAG 1800
1741 CGGAGTGAATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGCTGCTGAAGAAATTCAG 1800
1801 CGGAAAGCAGAGATGTCAGCAATACTGTTGAAGCTTAAGATCAAGGAGAGCTGAA 1860
1801 CGGAAAGCAGAGATGTCAGCAATACTGTTGAAGCTTAAGATCAAGGAGAGCTGAA 1860
1861 GTGGGAGATATCGGAAATCGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
1861 GTGGGAGATATCGGAAATCGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
1921 CTCAGAGAGAACTGGAGAAAGCT----- 1944
1921 CTCAGAGAGAACTGGAGAAAGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGTGTCAG 1980
1945 -----CAGAGAGCAGCCGAGAGGAGCTCGAAGAGCTGAGAGACCGAGAG 1992
1981 AATATCGGCAGCAAGAGCGAGCGAGGAGCTCGAGAGCTCGAGAGCTCGAGAACCGAGAG 2040
1993 GATTCCTTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAGCGCCGCAATCTCTG 2052
2041 GATTCCTTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAGCGCCGCAATCTCTG 2100
2053 GAGAAACAAGTAAAGAGCTAGAGACCATGGAGCGTGAAGAAACAGACTGAAGATGAC 2112
2101 GAGAAACAAGTAAAGAGCTAGAGACCATGGAGCGTGAAGAAACAGACTGAAGATGAC 2160
2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGAGCTCGAAGAG 2172
2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGAGCTCGAAGAG 2220
2173 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTTAGAAGTGCACCTGAAACAGAAAGAG 2232
2221 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTTAGAAGTGCACCTGAAACAGAAAGAG 2280
2233 CAGCACTATGAGGAAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2292
2281 CAGCACTATGAGGAAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2340
2293 AAGGAGACATGAGGAAACATGATCGAGACACGAGGAGGAGGCCCATGAGAGGGCAAA 2352
2341 AAGGAGACATGAGGAAACATGATCGAGACACGAGGAGGAGGCCCATGAGAGGGCAAA 2400
2353 ATTCTCAGCGAAACAGAGCGGATGATCAATGCTATGATGATTCCAAGTCCAGATCCCTGGAA 2412
2401 ATTCTCAGCGAAACAGAGCGGATGATCAATGCTATGATGATTCCAAGTCCAGATCCCTGGAA 2460
2413 CAGAGGATGTGAACTGTGAAGCCAAATAAATTCGAGCAAAATAGCAGTCTTTTACC 2472
2461 CAGAGGATGTGAACTGTGAAGCCCAAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
2473 CAAGGAAACATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
2521 CAAGGAAACATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAACTGGAGGAGCAGCTGGAG 2592

2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2640
2593 AAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGCTGGAACTGGAGACAAAGATTG 2652
2641 AAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGCTGGAACTGGAGACAAAGATTG 2700
2653 CGGAGGTGATCTATAGAGCAGCAGGAGCAGAACTGGAGCTCAAGCCGAGCTCAGAGAG 2712
2701 CGGAGGTGATCTATAGAGCAGCAGGAGCAGAACTGGAGCTCAGCCGAGCTCAGAGAG 2760
2713 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGACAGGCTGACAGCGCG 2772
2761 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGACAGGCTGACAGCGCG 2820
2773 GCGCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCA CAGCAGAAGCT 2832
2821 GCGCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCA CAGCAGAAGCT 2880
2833 GAAGAGGAGATCCAGGCACTCAAGGCACTAGAGATGAATCCAGCGCAATTTGATGCT 2892
2881 GAAGAGGAGATCCAGGCACTCAAGGCACTAGAGATGAATCCAGCGCAATTTGATGCT 2940
2893 CTTCTAACAGCTGATCTGTAATCAGAGCTGGAGGAGCAGCTAAACACAGCTGACCGAG 2952
2941 CTTCTAACAGCTGATCTGTAATCAGAGCTGGAGGAGCAGCTAAACACAGCTGACCGAG 3000
2953 GACAACTGTAATCAACAAACAACTTCTACTTGTCTCAAAACAACTCGATGAGGCTTCT 3012
3001 GACAACTGTAATCAACAAACAACTTCTACTTGTCTCAAAACAACTCGATGAGGCTTCT 3060
3013 GCGCCCAACAGCAGATGTTGTAACCTGCGAAGTGAAGTGACCATCTCCGCCGAGGAGATC 3072
3061 GCGCCCAACAGCAGATGTTGTAACCTGCGAAGTGAAGTGACCATCTCCGCCGAGGAGATC 3120
3073 ACGAAACAGAGATGAGCTTTACAGCCAGAAACAGATGAGGCTCTGAAGACACAG 3132
3121 ACGAAACAGAGATGAGCTTTACAGCCAGAAACAGATGAGGCTCTGAAGACACAG 3180
3133 TGCACTGCTGGAGGAAACAGGCTCATGATTTGAGGCCCTTAACATGAGCTGTAGAA 3192
3181 TGCACTGCTGGAGGAAACAGGCTCATGATTTGAGGCCCTTAACATGAGCTGTAGAA 3240
3193 AAAGAGCGCAGTGGAGGCTTGAGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTTGA 3252
3241 AAAGAGCGCAGTGGAGGCTTGAGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTTGA 3300
3253 TGTCGGGTTGAGAGCTGAGAGATGCTGACACCGGAAACAGAGCAGGCGGAGAGCC 3312
3301 TGTCGGGTTGAGAGCTGAGAGATGCTGACACCGGAAACAGAGCAGGCGGAGAGCC 3360
3313 GATCAGCGGATCACCGAGTCTCGCCAGGTTGGAGCTGCGAGTGAAGGAGCAAGAGCT 3372
3361 GATCAGCGGATCACCGAGTCTCGCCAGGTTGGAGCTGCGAGTGAAGGAGCAAGAGCT 3420
3373 GAGATTCCTGCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCTCTCT 3432
3421 GAGATTCCTGCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCTCTCT 3480
3433 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCCGAGCTTA 3492
3481 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCCGAGCTTA 3540
3493 CAGCAGAAGCTGAGACTGAAGCAGAGCTCAAAACAGAGGCTCTGGAAGAGCAAGCCAAA 3552
3541 CAGCAGAAGCTGAGACTGAAGCAGAGCTCAAAACAGAGGCTCTGGAAGAGCAAGCCAAA 3600
3553 TTACAGCAGAGATGAGCTGAGAGAAATCAATTTTCCGCTGACTCAAGGAGCTGCAA 3612
3601 TTACAGCAGAGATGAGCTGAGAGAAATCAATTTTCCGCTGACTCAAGGAGCTGCAA 3660
3613 GAAAGCTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3672
3661 GAAAGCTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3720

QY 3673 GAAAAATTCAGTCTTCTATTTCTCATGAAAAGGTGAAAAATGGAAGGCACTATTCTCAA 3732
Db 3721 GAAAAATTCAGTCTTCTATTTCTCATGAAAAGGTGAAAAATGGAAGGCACTATTCTCAA 3780
QY 3733 CAAACCAAATCATTTGATTTCTGCAAGCCAAATGGAACCACTGCTTAAAAAGAAAAG 3792
Db 3781 CAAACCAAATCATTTGATTTCTGCAAGCCAAATGGAACCACTGCTTAAAAAGAAA-- 3838
QY 3793 GGTTTATTAGTCGACGGAAAGAGGACCCTGCTTTACCCACACAGGTTTCTCTGCACTAC 3852
Db 3839 -----AGTTTCTCTGCACTAC 3855
QY 3853 AATGAGCTGAAGCTGGCCCTGAGAAAGAGAAAGCTGCTGTGCAAGCTAGAGAAAGCC 3912
Db 3856 AATGAGCTGAAGCTGGCCCTGAGAAAGAGAAAGCTGCTGTGCAAGCTAGAGAAAGCC 3915
QY 3913 CTTTCAAGAGACCGCATCGAGCTCCGCTCCGCCCGGAGGAGCTGCCACCGCAAGCA 3972
Db 3916 CTTTCAAGAGACCGCATCGAGCTCCGCTCCGCCCGGAGGAGCTGCCACCGCAAGCA 3975
QY 3973 ACGGACCAACCCATCCACGCCAGCCACCGCGAGGCAAGCATGCCATGTCGCC 4032
Db 3976 ACGGACCAACCCATCCACGCCAGCCACCGCGAGGCAAGCATGCCATGTCGCC 4035
QY 4033 ATCGTGGGTGCGAGAGCAGCCAGTCCAGTSCATGAGCTGCTGGCCCGCCATCCAGC 4092
Db 4036 ATCGTGGGTGCGAGAGCAGCCAGTCCAGTSCATGAGCTGCTGGCCCGCCATCCAGC 4095
QY 4093 CGCAGAAAGGAGCTTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAACGCATGCAC 4152
Db 4096 CGCAGAAAGGAGCTTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAACGCATGCAC 4155
QY 4153 CACAATATTCCTCAACGGATTCAACGTAGGACTGGAACATCGAGAGCCACAAAGTGTGCTGT 4212
Db 4156 CACAATATTCCTCAACGGATTCAACGTAGGACTGGAACATCGAGAGCCACAAAGTGTGCTGT 4215
QY 4213 TGTCTGGATACCGTGCACCTTGGAGCGCCAGGCAATCCAAATGTCGAATGTCAGTGTATG 4272
Db 4216 TGTCTGGATACCGTGCACCTTGGAGCGCCAGGCAATCCAAATGTCGAATGTCAGTGTATG 4275
QY 4273 TGTACCCCAAGTGTCCACGTGCTTGCAGCCACCTGGCGTTCCCTGCTGAATATGCC 4332
Db 4276 TGTACCCCAAGTGTCCACGTGCTTGCAGCCACCTGGCGTTCCCTGCTGAATATGCC 4335
QY 4333 ACACATTCACCGAGCCTTCTCCGTGACAAATGAACTCCCGAGTCTCCAGACCAAG 4392
Db 4336 ACACATTCACCGAGCCTTCTCCGTGACAAATGAACTCCCGAGTCTCCAGACCAAG 4395
QY 4393 GAGCCAGCAGCAGCTTGCACTGGAAGGGTGGATGAAGGTGCCAGGAATAACAAACGA 4452
Db 4396 GAGCCAGCAGCAGCTTGCACTGGAAGGGTGGATGAAGGTGCCAGGAATAACAAACGA 4455
QY 4453 GGACAGCAAGGTGGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCTATTTAT 4512
Db 4456 GGACAGCAAGGTGGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCTATTTAT 4515
QY 4513 GACAAATGAAGCCAGAGAACTGACAGAGCCGGTGGAGAAATTTGAGCTGTCCTTCCC 4572
Db 4516 GACAAATGAAGCCAGAGAACTGACAGAGCCGGTGGAGAAATTTGAGCTGTCCTTCCC 4575
QY 4573 GACGGGATGTATCTATTATGTTGCTGCTGCTTCCGAACTCCGAAATACGCCAAA 4632
Db 4576 GACGGGATGTATCTATTATGTTGCTGCTGCTTCCGAACTCCGAAATACGCCAAA 4635
QY 4633 GCA----- 4635
Db 4636 GCAGATGTCCCATACATATGAAGATGGAAATCTCACCCGCAACCACTGTGGCCCGGG 4695
QY 4636 ----- 4635
Db 4696 AGAACCTCTACTGTAGCTCCAGCTTCCCTTGACAAACAGCGCTGGGTCCACGCCCTTA 4755

QY 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4562
Db 4756 GAATCAGTTTGTGCAAGTGGAGAGTTTCTAGGAAAAAGCAGAAAGCTGATGCTAAACTG 4815
QY 4663 CTTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGCTAGACATGAACCTGCAACGCTG 4722
Db 4816 CTTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGCTAGACATGAACCTGCAACGCTG 4875
QY 4723 CCTTTCAGTACACAGGTGCTGTTGGTGGGCAACGAGAAAGGGCTCTACGCCCTGAAATGTC 4782
Db 4876 CCTTTCAGTACACAGGTGCTGTTGGTGGGCAACGAGAAAGGGCTCTACGCCCTGAAATGTC 4935
QY 4783 TTGAAAAACTCCCTTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAAATTTATATTATC 4842
Db 4936 TTGAAAAACTCCCTTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAAATTTATATTATC 4995
QY 4843 AAGGACCTGTGAGAAAGCTACTCATATAGCAGGAGAAAGAGCGGCACCTGTCTCTTGTGGAC 4902
Db 4996 AAGGACCTGTGAGAAAGCTACTCATATAGCAGGAGAAAGAGCGGCACCTGTCTCTTGTGGAC 5055
QY 4903 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACACATCTCA 4962
Db 5056 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACACATCTCA 5115
QY 4963 CCACAATTTTGAAGCTGTCAAAGGGCTGCCACTTGTTTGGGCGAGGCAAGATTGAGAAC 5022
Db 5116 CCACAATTTTGAAGCTGTCAAAGGGCTGCCACTTGTTTGGGCGAGGCAAGATTGAGAAC 5175
QY 5023 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTGCTATTTCTCCGCTACAAAGAAAAC 5082
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTGCTATTTCTCCGCTACAAAGAAAAC 5235
QY 5083 CTGAGCAAAATGCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5142
Db 5236 CTGAGCAAAATGCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5295
QY 5143 TTCAACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTTACGAAATTCGATGAAGCAG 5202
Db 5296 TTCAACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTTACGAAATTCGATGAAGCAG 5355
QY 5203 TACAGCTCGAGGAATTCCTGATAGAAATGACCAATTCCTTGGACCTGCTGTTGTTGCC 5262
Db 5356 TACAGCTCGAGGAATTCCTGATAGAAATGACCAATTCCTTGGACCTGCTGTTGTTGCC 5415
QY 5363 GCCTTTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACAGCGCAGGCGACGAGAG 5322
Db 5416 GCCTTTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACAGCGCAGGCGACGAGAG 5475
QY 5323 GAGTACTTGTGCTGTTTCCACGAAATTTGGAGTGTTCGTGGAATTTCTTACGGAAGACGTAGC 5382
Db 5476 GAGTACTTGTGCTGTTTCCACGAAATTTGGAGTGTTCGTGGAATTTCTTACGGAAGACGTAGC 5535
QY 5383 CGCACACAGCATCTCAAGTGGAGTTCCTTACCTTTGGCCCTTTGCCTACAGAGAACCCCTAT 5442
Db 5536 CGCACACAGCATCTCAAGTGGAGTTCCTTACCTTTGGCCCTTTGCCTACAGAGAACCCCTAT 5595
QY 5443 CTGTTTGTGACCACTTCAAATCACTCGAAGTAAATTTGAGATTCGAGGCAACGCTCCTCAGCA 5502
Db 5596 CTGTTTGTGACCACTTCAAATCACTCGAAGTAAATTTGAGATTCGAGGCAACGCTCCTCAGCA 5655
QY 5503 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5562
Db 5656 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5715
QY 5563 TCCTCAGAGCGATTTTACTTTGGGCTCCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5622
Db 5716 TCCTCAGAGCGATTTTACTTTGGGCTCCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5775
QY 5623 AAGGGAACCTCGTGAAGGAGTCCGGACATGAAACACCCGGGCGCGTCCACCTCCCGC 5682
Db 5776 AAGGGAACCTCGTGAAGGAGTCCGGACATGAAACACCCGGGCGCGTCCACCTCCCGC 5835
QY 5683 AGCAGCCCAACAGCGAGGCGCCACCCACCTGACAAAGAGCACATCAACAAAGCGGTGCCC 5742

Db 5836 AGCAGCCCAACAGCGAGCGCCACCACTGACACGAGCACATCATCAGCGCGTGGCC 5895
Qy 5743 TCAGCCACAGCGCGCGCGAGCGCCAGCCACCGCGAGAGCGCAACGCCACCCACCGC 5802
Db 5896 TCAGCCACAGCGCGCGCGAGCGCCAGCCACCGCGAGAGCGCAACGCCACCCACCGC 5955
Qy 5803 TACCGAGCGCGCGCGAGCGCGAGCGCAACGTCCTGCGCGCGCGCGCGCGCGCGCG 5862
Db 5956 TACCGAGCGCGCGCGAGCGCGAGCGCGCAACGTCCTGCGCGCGCGCGCGCGCGCG 6015
Qy 5863 GAGAGTCCCGCGCGCGAGTGTCTAGCAGCGCGAGAGCGGTCCCGCGGAGGCTGTTT 5922
Db 6016 GAGAGTCCCGCGCGCGAGTGTCTAGCAGCGCGAGAGCGGTCCCGCGGAGGCTGTTT 6075
Qy 5923 GAAGACAGCAGCGCGCGCGCGTGTGCGGAGCGCGTGTGCGGAGCGCGTGTGCGGAG 5982
Db 6076 GAAGACAGCAGCGCGCGCGCGTGTGCGGAGCGCGTGTGCGGAGCGCGTGTGCGGAG 6135
Qy 5983 AACAGGGAAGAGCGCGAGGTGC 6005
Db 6136 AACAGGTCGTGGACCACTTTC 6158

RESULT 7

US-10-262-511-1
; Sequence 1, Application US/10262511
; Publication No. US20040038223A1

GENERAL INFORMATION:

; APPLICANT: Smithson, Glenda
; APPLICANT: Mallet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eilerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Blina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curasequid version 0.1
; SEQ ID NO 1
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6159)
US-10-262-511-1
Query Match 91.4%; Score 5631; DB 13; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;
Qy 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGTGTAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGTGTAACCCATT 60
Qy 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACCCCTTTATGACTCAA 120
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACCCCTTTATGACTCAA 120
Qy 121 CAGCAGATGTCTCCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGTCTCCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTTTGAA 180
Qy 181 GAATGCACTGAGCTGCTCTGTATGAAGATTAAAGCACTGTGTCGGAAGTAT 240
Db 181 GAATGCACTGAGCTGCTCTGTATGAAGATTAAAGCACTGTGTCGGAAGTAT 240
Qy 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAGAGCTTCGGAAGTCA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAGAGCTTCGGAAGTCA 300
Qy 301 AGTCTTCTAGGTTGTGTCACCTTTGCTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
Db 301 AGTCTTCTAGGTTGTGTCACCTTTGCTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
Qy 361 GACATCTATGCTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
Db 361 GACATCTATGCTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
Qy 421 TTTTTCGAGGAGCGGAAACATATTATCTCGAAGCACAAAGCCCGTGGATTCGCCAATTA 480
Db 421 TTTTTCGAGGAGCGGAAACATATTATCTCGAAGCACAAAGCCCGTGGATTCGCCAATTA 480
Qy 481 CAGTATGCTTTCAGGCAAAATACCTTTATCTGATGAGGAATATCAGCTCGAGGG 540
Db 481 CAGTATGCTTTCAGGCAAAATACCTTTATCTGATGAGGAATATCAGCTCGAGGG 540
Qy 541 GACTTGTCTGCTACCTTTTGAATAGATATGAGGACAGTTAGATGAAACCTGTACAGTTT 600
Db 541 GACTTGTCTGCTACCTTTTGAATAGATATGAGGACAGTTAGATGAAACCTGTACAGTTT 600
Qy 601 TACCTAGCTGAGCTGATTTTGGCTGTTCACAGGCTTCATCTGATGAGTACGTCATCA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTCACAGGCTTCATCTGATGAGTACGTCATCA 660
Qy 661 GACATCAAGCTGAGAACATCTTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
Db 661 GACATCAAGCTGAGAACATCTTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720

QY 721 GGATCTGCGCGAAATGAATTCRAACAGAGATGGTGAATGCCAACTCCCGATTGGGACC 780
Db 721 GGATCTGCGCGAAATGAATTCRAACAA ---GGTGAATGCCAACTCCCGATTGGGACC 777
QY 781 CCAGATTACATGGCTCCCTCAAGTGTGACTGTGATGAACGGGGATGGAAAGGCACCTAC 840
Db 778 CCAGATTACATGGCTCCCTCAAGTGTGACTGTGATGAACGGGGATGGAAAGGCACCTAC 837
QY 841 GSCCTGGACTGTGACTGGTGGTCACTGGGCGTGATGCTTATGAGATGATTTATGGGAGA 900
Db 838 GSCCTGGACTGTGACTGGTGGTCACTGGGCGTGATGCTTATGAGATGATTTATGGGAGA 897
QY 901 TCCCTCTTCGACAGAGGAACTCTGCAGAACCTTCAATAACATTATGAATTTCCAGCGG 960
Db 898 TCCCTCTTCGACAGAGGAACTCTGCAGAACCTTCAATAACATTATGAATTTCCAGCGG 957
QY 961 TTTTGAATTCAGATGACCCCAAGTGACAGTACTTTCTTGTATCTGATCTGATCAAAAGC 1020
Db 958 TTTTGAATTCAGATGACCCCAAGTGACAGTACTTTCTTGTATCTGATCTGATCAAAAGC 1017
QY 1021 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCGCATCTTCTTC 1080
Db 1018 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCGCATCTTCTTC 1077
QY 1081 TCTAAATTTGACTGGAAACAATTCGTAACTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1140
Db 1078 TCTAAATTTGACTGGAAACAATTCGTAACTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1137
QY 1141 TCTGAGATGACACCTCCAAATTTTGTATGAACAGAGAGAAATTCGTGGGTTTCATCCTCT 1200
Db 1138 TCTGAGATGACACCTCCAAATTTTGTATGAACAGAGAGAAATTCGTGGGTTTCATCCTCT 1197
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTCCCGTTTGTGGGTTTTTCG 1260
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTCCCGTTTGTGGGTTTTTCG 1257
QY 1261 TACAGCAAGGACTGGGATTCCTGGTAGATCTGAGTCTGTGTCTGGGCTTGAGCTCC 1320
Db 1258 TACAGCAAGGACTGGGATTCCTGGTAGATCTGAGTCTGTGTCTGGGCTTGAGCTCC 1317
QY 1321 CTTGCCAAGCTAGCTCCATGAAAGAACTTCATCAAAAGCAAAAGAGCTCAAGAC 1380
Db 1318 CTTGCCAAGCTAGCTCCATGAAAGAACTTCATCAAAAGCAAAAGAGCTCAAGAC 1377
QY 1381 TCTCAGGACAAGTGTCAAGAATGGAGCAGGAAATGACCCGTTTACATCGGAGAGTGTCA 1440
Db 1378 TCTCAGGACAAGTGTCAAGAATGGAGCAGGAAATGACCCGTTTACATCGGAGAGTGTCA 1437
QY 1441 GAGGTGGAGCTGTGCTTAGTCAGAAGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500
Db 1438 GAGGTGGAGCTGTGCTTAGTCAGAAGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1497
QY 1501 TCCCTCTCGAGCAGGACCTTGCTACTACATCAAGAAATGACGTAGCTTAAAGCGAAAGT 1560
Db 1498 TCCCTCTCGAGCAGGACCTTGCTACTACATCAAGAAATGACGTAGCTTAAAGCGAAAGT 1557
QY 1561 TTGGAGCAAGCAGATGAGGTGTCCAGAGGAGATGACAAAGCACTGAGCTTCTCCAT 1620
Db 1558 TTGGAGCAAGCAGATGAGGTGTCCAGAGGAGATGACAAAGCACTGAGCTTCTCCAT 1617
QY 1621 GATATCAGAGAGCAGAGCGGAGCTTCCAAAGAAATCAAAAGCAGAGGTACCAAGCTCAA 1680
Db 1618 GATATCAGAGAGCAGAGCGGAGCTTCCAAAGAAATCAAAAGCAGAGGTACCAAGCTCAA 1677
QY 1681 GTGGAGAAATGAGGTGTGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740
Db 1678 GTGGAGAAATGAGGTGTGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAGA 1737
QY 1741 CGAGTGTATCTTACCAATCTGAGCTGAGAGGTCTCGGCTTCTGCTGAAGAAATCAAG 1800
Db 1738 CGAGTGTATCTTACCAATCTGAGCTGAGAGGTCTCGGCTTCTGCTGAAGAAATCAAG 1797
QY 1801 CGAAAGCGCAGAAATGTCAAGCAATACTGTTGAAGGCTTAAGGATCAAGGGAAAGCTGAA 1860

Db 1798 CGAAAGCGCAGAAATGTCAAGCATAACTGTTGAAGCTAAGGATCAGGGGAGGCTGAA 1857
QY 1861 GTGGAGAAATATCGGAAACTGGAGAGATCAATGCTTGAGCAGCAGCTCAAAATTCAGGAG 1920
Db 1858 GTGGAGAAATATCGGAAACTGGAGAGATCAATGCTTGAGCAGCAGCTCAAAATTCAGGAG 1917
QY 1921 CTCACAGAGAAACTGGAGAAAGGCTG----- 1945
Db 1918 CTCACAGAGAAACTGGAGAAAGGCTGTTAAAGCCAGCACCGAGGCCACCGAGCTGCTGCAG 1977
QY 1946 -----CAAGAGCGAGCGAGAGGAGCTGGAGAGCTGGAGAACTGCAGAAACCGAGAG 1992
Db 1978 AATATCCCGCAGGCAMAGGAGCGAGGAGGAGCTGGAGAGGCTGCAGAAACCGAGAG 2037
QY 1993 GATTCTTCTGAAGCATCAGAAAGAGCTGGTGGAGCTGAGGAAACGCGCCCATCTCTCTG 2052
Db 2038 GATTCTTCTGAAGCATCAGAAAGAGCTGGTGGAGCTGAGGAAACGCGCCCATCTCTCTG 2097
QY 2053 GAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAAACAGACTGAAGGATGAC 2112
Db 2098 GAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAAACAGACTGAAGGATGAC 2157
QY 2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAATCTGAGAGCTCGAGAG 2172
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAATCTGAGAGCTCGAGAG 2217
QY 2173 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2232
Db 2218 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2277
QY 2233 CAGCACTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAGAAAGACCTGGCTGAC 2292
Db 2278 CAGCACTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAGAAAGACCTGGCTGAC 2337
QY 2293 AAGGAGACCTGGAGAACATGATGCAGACACGAGGAGGAGGCCCATGAGAGGGCAAA 2352
Db 2338 AAGGAGACCTGGAGAACATGATGCAGACACGAGGAGGAGGCCCATGAGAGGGCAAA 2397
QY 2353 ATCTCAGCAACAGAAAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGAA 2412
Db 2398 ATCTCAGCAACAGAAAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGAA 2457
QY 2413 CAGAGGATTTGAGACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGCAGCTCTTTTACC 2472
Db 2458 CAGAGGATTTGAGACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGCAGCTCTTTTACC 2517
QY 2473 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
Db 2518 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577
QY 2533 CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAACCCGAAAACTGGAGGAGCAGCTGGAG 2592
Db 2578 CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAACCCGAAAACTGGAGGAGCAGCTGGAG 2637
QY 2593 AAGATCAGCCACCAAGACCACTGCAAGAANTCGGCTGCTGGAACCTGGAGACAGAAATTG 2652
Db 2638 AAGATCAGCCACCAAGACCACTGCAAGAANTCGGCTGCTGGAACCTGGAGACAGAAATTG 2697
QY 2653 CGGAGGCTCAGTCTTAGAGCACGAGGAGCAGAAACTGGAGCTCAAGCCCGCAGCTCAGAGAG 2712
Db 2698 CGGAGGCTCAGTCTTAGAGCACGAGGAGCAGAAACTGGAGCTCAAGCCCGCAGCTCAGAGAG 2757
QY 2713 CTACAGCTCTCCCTGAGGAGCGGAGTCACTGTTGACGCTTGCAGCTGCGAGCTGCACGGGCG 2772
Db 2758 CTACAGCTCTCCCTGAGGAGCGGAGTCACTGTTGACGCTTGCAGCTGCGAGCTGCACGGGCG 2817
QY 2773 GCCCTGAGAGCCAGCTTCGCCAGGCCAGACAGAGCTGGAAAGAGACCAAGAGAGCT 2832
Db 2818 GCCCTGAGAGCCAGCTTCGCCAGGCCAGACAGAGCTGGAAAGAGACCAAGAGAGCT 2877
QY 2833 GAAGAGAGATTCAGGCACTCAAGGCAATAGAGATGAAATCCAGCCCAATTTGATGCT 2892

QY 4393 GAGCCAGCAGCAGCTTGCACCTGGAAGGCTGATGAAGTGCCTCCAGGAATAACAACGA 4452
Db 4393 GAGCCAGCAGCAGCTTGCACCTGGAAGGCTGATGAAGTGCCTCCAGGAATAACAACGA 4452
QY 4453 GGACAGCAAGCTGGGACAGAGATCAATTTCTTGGAGGATCAAAAGTCCTCATTTAT 4512
Db 4453 GGACAGCAAGCTGGGACAGAGATCAATTTCTTGGAGGATCAAAAGTCCTCATTTAT 4512
QY 4513 GACAAATCAAGCAGAGAGCTGGACAGAGCCGCTGGAAGATTTGAGCTGTGCTTCCC 4572
Db 4513 GACAAATCAAGCAGAGAGCTGGACAGAGCCGCTGGAAGATTTGAGCTGTGCTTCCC 4572
QY 4573 GAGGGGATGATATCTATTCATGCTGGCTTGGCTTCCGAAATCGCAAAATACAGCCAAA 4632
Db 4573 GAGGGGATGATATCTATTCATGCTGGCTTGGCTTCCGAAATCGCAAAATACAGCCAAA 4632
QY 4633 GCA----- 4635
Db 4633 GCAGATGCCATACATATCTGAAGATGGAATCTCACCCGACACACCTGCTGGCCCGGG 4692
QY 4636 ----- 4635
Db 4693 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTACCGGCTTA 4752
QY 4636 -----GAAAAAGCAGAGCTGATGCTTAACTG 4662
Db 4753 GAATCAGTTCTCCAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTTAACTG 4812
QY 4663 CTGGAACCTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4722
Db 4813 CTGGAACCTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4872
QY 4723 CCTTCAGTGACAGGTGTTGTTGGGACCGAGGAAGGGCTCTACCCCTGAATGTC 4782
Db 4873 CCTTCAGTGACAGGTGTTGTTGGGACCGAGGAAGGGCTCTACCCCTGAATGTC 4932
QY 4783 TTGAAAACTCCCTAACCCATGTCACAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4842
Db 4933 TTGAAAACTCCCTAACCCATGTCACAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4992
QY 4843 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGCTTGTGGAC 4902
Db 4993 AAGGACCTGGAGAGCTACTCATGATAGCAGGTAAGAGCGGCACTGTGCTTGTGGAC 5052
QY 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4962
Db 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5112
QY 4963 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5022
Db 5113 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5172
QY 5023 GGGCTCTGATCTGTGACGCGATGCCAGCAAGTGTCAATTTCTCGCTACAAAGAAAC 5082
Db 5173 GGGCTCTGATCTGTGACGCGATGCCAGCAAGTGTCAATTTCTCGCTACAAAGAAAC 5232
QY 5083 CTCAGCAATATCTGCATCCGAAAGAGATAGAGCTCAGAGCCCTGAGCTGTATCCAC 5142
Db 5233 CTCAGCAATATCTGCATCCGAAAGAGATAGAGCTCAGAGCCCTGAGCTGTATCCAC 5292
QY 5143 TTACCAATATCAGTATCTCTATTTGAACCAATAATTTACGAATCGACATGAAGCAG 5202
Db 5293 TTACCAATATCAGTATCTCTATTTGAACCAATAATTTACGAATCGACATGAAGCAG 5352
QY 5203 TACAGCTCGAGGAATTCCTGATAGATGACCATCTCTGGACCTGCTGTGTTGCC 5262
Db 5353 TACAGCTCGAGGAATTCCTGATAGATGACCATCTCTGGACCTGCTGTGTTGCC 5412
QY 5263 GCCTCTTCCACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGAGGAG 5322
Db 5413 GCCTCTTCCACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGAGGAG 5472
QY 5323 GAGTACTTGTCTGTTTCCAGAAATTTGGAGTGTGCTGGATTTCTACGGAGAGCTAGC 5382

Db 5473 GAGTACTTGTCTGTTTCCAGAAATTTGGAGTGTGCTGGATTTCTACGGAAGAGCTAGC 5532
QY 5383 GCACAGACGATCTCAAGTGGAGTGCCTTACTTTGGCTTTGCTACAGAGAACCCTAT 5442
Db 5533 GCACAGACGATCTCAAGTGGAGTGCCTTACTTTGGCTTTGCTACAGAGAACCCTAT 5592
QY 5443 CTGTTTGGAGCCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5502
Db 5593 CTGTTTGGAGCCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5652
QY 5503 GGGACCCCTGCCGAGCGTACTTGGACATCCGAAACCCGCTACCTTGGGCCCTTGCATT 5562
Db 5653 GGGACCCCTGCCGAGCGTACTTGGACATCCGAAACCCGCTACCTTGGGCCCTTGCATT 5712
QY 5563 TCCTCAGAGCGATTTACTTTGGCTCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5622
Db 5713 TCCTCAGAGCGATTTACTTTGGCTCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5772
QY 5623 AAGGAAACCTCGTGAAGGAGTCCGCACTGAAACACCCGCGGCCCTCCACCTCCGC 5682
Db 5773 AAGGAAACCTCGTGAAGGAGTCCGCACTGAAACACCCGCGGCCCTCCACCTCCGC 5832
QY 5683 AGCAGCCCAACAGAGAGCGCCACCACTGTAACAGACATCACAAAGCGGTGGCC 5742
Db 5833 AGCAGCCCAACAGAGAGCGCCACCACTGTAACAGACATCACAAAGCGGTGGCC 5892
QY 5743 TCCAGCCAGCGCCGCGGAGGCGCCAGCACCCGCGAGAGCAAGCAACACCCACCGC 5802
Db 5893 TCCAGCCAGCGCCGCGGAGGCGCCAGCACCCGCGAGAGCAAGCAACACCCACCGC 5952
QY 5803 TACCGAGGCGGAGCCGAGCTGCGCAGGGAAGTCTCTGCGCCGCCCTTGGAGCGA 5862
Db 5953 TACCGAGGCGGAGCCGAGCTGCGCAGGGAAGTCTCTGCGCCGCCCTTGGAGCGA 6012
QY 5863 GAGAAGTCCCGCGCGGATGCTCAGCACCGGAGAGCGGTCCCGCGGAGGCTGTTT 5922
Db 6013 GAGAAGTCCCGCGCGGATGCTCAGCACCGGAGAGCGGTCCCGCGGAGGCTGTTT 6072
QY 5923 GAAGACAGCAGCGGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGAGTG 5982
Db 6073 GAAGACAGCAGCGGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGAGTG 6132
QY 5983 AACAGGAGAGGCGAGAGTGCCTCTCAGTCTTACGTTTACACTGTCACT 6037
Db 6133 AACAGGAGAGGCGAGAGTGCCTCTCAGTCTTACGTTTACACTGTCACT 6187

RESULT 9

US-10-262-511-13

; Sequence 13, Application US/10262511

; Publication No. US2004003823A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glennda

; APPLICANT: Millet, Isabelle

; APPLICANT: Peyman, John A.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Ju, Jingfang

; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Patturajan, Meera

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Elletman, Karen

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Ort, Tatiana

; APPLICANT: Gorman, Linda

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Anderson, David W.

; APPLICANT: Zhong, Mei

; APPLICANT: Catterton, Elina

; APPLICANT: Ji, Weizhen

; APPLICANT: Miller, Charles E.


```

; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 13
; LENGTH: 6201
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6198)
US-10-262-511-13

Query March
Best Local Similarity 91.4%; Score 5629.8; DB 13; Length 6201;
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;

QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT 60
QY 61 GCAGCGGGCTCCAGGCTGATCTGTTCCAGGGGAACCCACCTTTATGACTCAA 120
Db 61 GCAGCGGGCTCCAGGCTGATCTGTTCCAGGGGAACCCACCTTTATGACTCAA 120
QY 121 CAGCAGATGTCCTCTCTTTCCCGAGAGGATATAGATCCCTCTTTCTTTTGAA 180
Db 121 CAGCAGATGTCCTCTCTTTCCCGAGAGGATATAGATGCTCTTTCTTTTGAA 180
QY 181 GAATGAGTCCAGCTGCTCTGATGAAGATTAAGCAGTGAGCACTTCTCGGAGATAT 240
Db 181 GAATGAGTCCAGCTGCTCTGATGAAGATTAAGCAGTGAGCACTTCTCGGAGATAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGGACTTCGAAGTCAG 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGGACTTCGAAGTCAG 300
QY 301 AGTCTTGAGTTGTGGTCACTTTGTGTAAGTGCAGGTGTTAGAGAGAAAGCAACCGGG 360
Db 301 AGTCTTGAGTTGTGGTCACTTTGTGTAAGTGCAGGTGTTAGAGAGAAAGCAACCGGG 360
```

```

361 GACATCTATGCTATGAAAGTGATGAAGAAAGAGGCTTTTATGGCCAGAGCAGGTTTCA 420
361 GACATCTATGCTATGAAAGTGATGAAGAAAGAGGCTTTTATGGCCAGAGCAGGTTTCA 420
421 TTTTGTGAGAAAGAGCGGAACATATATCTCGAAGCAGACCAAGCCCGTGGATCCCCCAATTA 480
421 TTTTGTGAGAAAGAGCGGAACATATATCTCGAAGCAGACCAAGCCCGTGGATCCCCCAATTA 480
481 CAGTATGCTCTTCAGGACCAAAATACCTTTATCTGATGAGGAATATCAGCCTCGAGGG 540
481 CAGTATGCTCTTCAGGACCAAAATACCTTTATCTGATGAGGAATATCAGCCTCGAGGG 540
541 GACTTGTCTGCTACTTTTGAATAGATAGAGGACCACTTAGATGAAACCTGTATACAGTTT 600
541 GACTTGTCTGCTACTTTTGAATAGATAGAGGACCACTTAGATGAAACCTGTATACAGTTT 600
601 TACTAGCTGAGCTGATTTTGGCTGTTACAGGGTTCATCTGATGGGATACGTGCAATCGA 660
601 TACTAGCTGAGCTGATTTTGGCTGTTACAGGGTTCATCTGATGGGATACGTGCAATCGG 660
661 GACATCAAGCCTGAGAACATTTCTGTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
661 GACATCAAGCCTGAGAACATTTCTGTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
721 GGATCTCCCGGAAATGAATTCAGAACAGATGCTGAATGCCAACTCCCGATTGGGACC 780
721 GGATCTCCCGGAAATGAATTCAGAACAGATGCTGAATGCCAACTCCCGATTGGGACC 777
781 CCAGATTACATGCTCCTGAAAGTCTGACTGTGATGAACCGGGGATGAAAGACACTAC 840
778 CCAGATTACATGCTCCTGAAAGTCTGACTGTGATGAACCGGGGATGAAAGACACTAC 837
841 GGCTGAGCTGTGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
838 GGCTGAGCTGTGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 897
901 TCCCTCTCCAGAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGG 960
898 TCCCTCTCCAGAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGG 957
961 TTTTGTGAAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
958 TTTTGTGAAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1017
1021 TTGTTGTGGGCGGCGAAGAGAGACTGAAGTTTGAAGTCTTGGCTGCTGCTGCTGCTGCTTC 1080
1018 TTGTTGTGGGCGGCGAAGAGAGACTGAAGTTTGAAGTCTTGGCTGCTGCTGCTGCTGCTTC 1077
1081 TCTAAATTTGACTGGAACAAACATTTGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
1078 TCTAAATTTGACTGGAACAAACATTTGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1137
1141 TCTCAAGATGACCTCCAAATTTTGAATGAACACAGAGAGAAATTCGTGGGTTTCACTCTCT 1200
1138 TCTCAAGATGACCTCCAAATTTTGAATGAACACAGAGAGAAATTCGTGGGTTTCACTCTCT 1197
1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGGTTTTTCG 1260
1198 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGGTTTTTCG 1257
1261 TACAGCAAGGCACTGGGGAATTTTGGTAGATCTGAGTCTGTTGTGGTCTGGGTTCTGACTCC 1320
1258 TACAGCAAGGCACTGGGGAATTTTGGTAGATCTGAGTCTGTTGTGGTCTGGGTTCTGACTCC 1317
1321 CTGCGCAAGTACTGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTCAAGAGAC 1380
1318 CTGCGCAAGTACTGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTCAAGAGAC 1377
1381 TCTCAGGACAAAGTGTCAAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTC 1440
1378 TCTCAGGACAAAGTGTCAAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTC 1437
1441 GAGGTGGAGGCTGTGCTTAGTCAAGAGGAGTGAAGGCTCTGAGACTCAGAGA 1500
```

[illegible]

Db	2518	CAAAGAA	CATGAAGGCCAAGAAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC	2577
Qy	2533	CTGGAGACA	CAGGCTGGGAAGCTTGGAGGCCAGAACCGAAAACTGGAGGAGCAGCTGGAG	2592
Db	2578	CTGGAGACA	CACAGGCTGGGAAGTTTGGAGGCCAGAACCGAAAACTGGAGGAGCAGCTGGAG	2637
Qy	2593	AGATTCAGCCA	CCAGACCCAGTGAACAAGAATCGGTCTGTGAACTGGAGACAAGATTG	2652
Db	2638	AAGATCAGCCA	CCAGACCCAGTGAACAAGAATCGGTCTGTGAACTGGAGACAAGATTG	2697
Qy	2653	CGGAGGTTCAGTCT	TAGAGCAGCAGGAGCAGAAATCGAGCTCAAGCGCCAGCTCACAGAG	2712
Db	2698	COGAGGTGAGTCT	TAGAGCAGCAGGAGCAGAAATCGAGCTCAAGCGCCAGCTCACAGAG	2757
Qy	2713	CTACAGCTCTCCCT	TCGAGGCGCGAGTTCACAGTTCGACAGCCCTGCAAGCTGCAACGGGG	2772
Db	2758	CTACAGCTCTCCCT	TCGAGGCGCGAGTTCACAGTTCGACAGCCCTGCAAGCTGCAACGGGG	2817
Qy	2773	GCCTGGAGAGC	AGAGCTTCGCCAGGCGAAAGACAGAGCTGGAAGAGACCAAGAGAGCT	2832
Db	2818	GCCTGGAGAGC	AGAGCTTCGCCAGGCGAAAGACAGAGCTGGAAGAGACCAAGAGAGCT	2877
Qy	2833	GAAGAGGAGATCCA	CGGCACTCAGGCGACATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2892
Db	2878	GAAGAGGAGATCCA	CGGCACTCAGGCGACATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2937
Qy	2893	CTTCGTTAACAGCT	TGATCTGTAATCAACAGACTGGAGAGCAGCTAAACAGCTGACCGAG	2952
Db	2938	CTTCGTTAACAGCT	TGATCTGTAATCAACAGACTGGAGAGCAGCTAAACAGCTGACCGAG	2997
Qy	2953	GACAACTGAACTCA	CAACACCAAACTTCTACTCTGCCAAACAACTCCATCAGCAGCTTCT	3012
Db	2998	GACAACTGCTGAA	CTCAACAAACCAAACTTCTACTCTGCCAAACAACTCCATCAGCAGCTTCT	3057
Qy	3013	GGGCCAAACGAC	AGATTTGTAACCTGCGAAGTGAAGTGGACCATCTCCGCGGGAGATC	3072
Db	3058	GGGCCAAACGAC	AGATTTGTAACCTGCGAAGTGAAGTGGACCATCTCCGCGGGAGATC	3117
Qy	3073	ACGGAACGAGAGAT	CAGCTTACACGCGAAGACCAACGATGGAGGCTCTGAAGACCAAG	3132
Db	3118	ACGGAACGAGAGAT	CAGCTTACACGCGAAGACCAACGATGGAGGCTCTGAAGACCAAG	3177
Qy	3133	TGCACCATGCTG	GAGAAACAGETCATGGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA	3192
Db	3178	TGCACCATGCTG	GAGAAACAGETCATGGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA	3237
Qy	3193	AAAGACGGCAGT	TGGAGAGGCTCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGAG	3252
Db	3238	AAAGACGGCAGT	TGGAGAGGCTCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGAG	3297
Qy	3253	TGTCCGGTTTCG	AGAGCTGCAGAGAAATGCTGGAACCCGAGAAACAGACAGGCGAGAGCC	3312
Db	3298	TGTCCGGTTTCG	AGAGCTGCAGAGAAATGCTGGAACCCGAGAAACAGACAGGCGAGAGCC	3357
Qy	3313	GATCAGGGATCAC	CGAGTCTCCGACAGTGTGGAGCTGGAGCTGGCAGTGAAGAGCACAAAGCT	3372
Db	3358	GATCAGGGATCAC	CGAGTCTCCGACAGTGTGGAGCTGGAGCTGGCAGTGAAGAGCACAAAGCT	3417
Qy	3373	GAGATTCTCGCT	CTGACAGAGCTCTCAAAGACAGAACTGAAAGGCCGAGAGCTCTCT	3432
Db	3418	GAGATTCTCGCT	CTGACAGAGCTCTCAAAGACAGAACTGAAAGGCCGAGAGCTCTCT	3477
Qy	3433	GACAACTCAAT	GACTGGAGAAAGCATGCTATGCTTGAATGAATGCCCGAGCTTA	3492
Db	3478	GACAACTCAAT	GACTGGAGAAAGCATGCTATGCTTGAATGAATGCCCGAGCTTA	3537
Qy	3493	CAGCAGACTGGAG	CTGAACGAGGCTCAACAGAGGCTCTTGGAGAGACAGACCAAA	3552
Db	3538	CAGCAGACTGGAG	CTGAACGAGGCTCAACAGAGGCTCTTGGAGAGACAGACCAAA	3597
Qy	3553	TTACAGCAGCAGAT	GGACCTGCAGAAAAATCACATTTTCCGCTCTGATCTCAAGGACTGCAA	3612
Db	3598	TTACAGCAGCAGAT	GGACCTGCAGAAAAATCACATTTTCCGCTCTGATCTCAAGGACTGCAA	3657

QY	3613	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATGGAGTATCAGCTG	3672	4636	-----	4635
Db	3658	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATGGAGTATCAGCTG	3717			
QY	3673	GAAGCAATTCAGAGTTCTCTATTCTCATGAAAGGTGAAATGGAGGCACTATTTCTCAA	3732			
Db	3718	GAAGCAATTCAGAGTTCTCTATTCTCATGAAAGGTGAAATGGAGGCACTATTTCTCAA	3777			
QY	3733	CAAAACCAAACTCATTTCTGCAAGCCAAATGACCACTGCTAAAGAAAAG	3792			
Db	3778	CAAAACCAAACTCATTTCTGCAAGCCAAATGACCACTGCTAAAGAAAAG	3835			
QY	3793	GGTTTATTAGTCGACGGAAGAGACCCCTCTTTAACCACACAGAGTTCTCTCAGTAC	3852			
Db	3836	-----AGGTGCGCTCTCAGTAC	3852			
QY	3853	AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGAGAGCTAGAGGAAGCC	3912			
Db	3853	AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGAGAGCTAGAGGAAGCC	3912			
QY	3913	CTTCAGAGACCCGATCGAGCTCGGTCCGCCCGGAGGAGCTGCCACCCGCAAGCA	3972			
Db	3913	CTTCAGAGACCCGATCGAGCTCGGTCCGCCCGGAGGAGCTGCCACCCGCAAGCA	3972			
QY	3973	ACGGACCAACCAACCCATCCACGCCAGCCACCGCGAGGAGAGATGCCATGTCCGCC	4032			
Db	3973	ACGGACCAACCAACCCATCCACGCCAGCCACCGCGAGGAGAGATGCCATGTCCGCC	4032			
QY	4033	ATCGTGGCGTGCAGAGACCAAGCCAGTGCATGAGCTGTGGCCCGGCCATCCAGC	4092			
Db	4033	ATCGTGGCGTGCAGAGACCAAGCCAGTGCATGAGCTGTGGCCCGGCCATCCAGC	4092			
QY	4093	CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAACGATGCCAC	4152			
Db	4093	CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAACGATGCCAC	4152			
QY	4153	CACAAATATTCCTCAACGATTAACCTAGGAATGAACTGAGAGCCACAAAGTGTGCTGTG	4212			
Db	4153	CACAAATATTCCTCAACGATTAACCTAGGAATGAACTGAGAGCCACAAAGTGTGCTGTG	4212			
QY	4213	TGCTGGATACCGTGCATCTTGAGCGCCAGGATCCAAATGTCTCGAATGTGAGTGATG	4272			
Db	4213	TGCTGGATACCGTGCATCTTGAGCGCCAGGATCCAAATGTCTCGAATGTGAGTGATG	4272			
QY	4273	TGTCACCCCAAGTGTCCACGTCCTTGCCAGCCACTCGGGCTTCGCTGTGAATATGCC	4332			
Db	4273	TGTCACCCCAAGTGTCCACGTCCTTGCCAGCCACTCGGGCTTCGCTGTGAATATGCC	4332			
QY	4333	ACACACTTCCAGAGCCCTTCTGCGTGAACAAATGAACTCCCGAGGCTCCAGACCAG	4392			
Db	4333	ACACACTTCCAGAGCCCTTCTGCGTGAACAAATGAACTCCCGAGGCTCCAGACCAG	4392			
QY	4393	GAGCCAGCAGAGCTTGCACTGGAAGGGTGGATGAAGGTGCCAGGAATAACAAACGA	4452			
Db	4393	GAGCCAGCAGAGCTTGCACTGGAAGGGTGGATGAAGGTGCCAGGAATAACAAACGA	4452			
QY	4453	GGACAGCAAGGTGGACAGGAAGTACATGTCTGAGAGGATCAAAAGTCTCTATTAT	4512			
Db	4453	GGACAGCAAGGTGGACAGGAAGTACATGTCTGAGAGGATCAAAAGTCTCTATTAT	4512			
QY	4513	GACAAATGAAGCCAGAGAGCTGACAGAGCGGTGGAAGAAATTTGAGCTGTGCCCTTCCC	4572			
Db	4513	GACAAATGAAGCCAGAGAGCTGACAGAGCGGTGGAAGAAATTTGAGCTGTGCCCTTCCC	4572			
QY	4573	GACGGGATGTATCTATTATCATGTGCGTGTGGTGTCTGCAACTGCGCAATACAGCCAAA	4632			
Db	4573	GACGGGATGTATCTATTATCATGTGCGTGTGGTGTCTGCAACTGCGCAATACAGCCAAA	4632			
QY	4633	GCA-----	4635			
Db	4633	GCAGATGTCCCATACATACATGAAGATGGAAATCTCACCCGACACACCTGCTGGCCCGGG	4692			

QY	4693	AGAACTCTACTTGTCTAGCTCCAGCTTCCCTGTGACAAAACAGCGTGGGTCAACCGCTTA	4752			
Db	4693	AGAACTCTACTTGTCTAGCTCCAGCTTCCCTGTGACAAAACAGCGTGGGTCAACCGCTTA	4752			
QY	4753	GAATCAGTTGTGCGAGTGGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATCTAACTG	4812			
Db	4753	GAATCAGTTGTGCGAGTGGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATCTAACTG	4812			
QY	4663	CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCCGCTTAGACATGAATCTGACGCTG	4722			
Db	4813	CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCCGCTTAGACATGAATCTGACGCTG	4872			
QY	4723	CCCTTCAGTGACCAAGTGTGTGTGGGACCGAGGAGGGCTCTACGCCCTGGAATGTC	4782			
Db	4873	CCCTTCAGTGACCAAGTGTGTGTGGGACCGAGGAGGGCTCTACGCCCTGGAATGTC	4932			
QY	4783	TTGAAAACTCCCTAAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC	4842			
Db	4933	TTGAAAACTCCCTAAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC	4992			
QY	4843	AAGGACCTGAGAGCTACTCATGATAGCAGAGAGAGCGGGCACTGTCTTGTGGAC	4902			
Db	4993	AAGGACCTGAGAGCTACTCATGATAGCAGAGAGAGCGGGCACTGTCTTGTGGAC	5052			
QY	4903	GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA	4962			
Db	5053	GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA	5112			
QY	4963	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCAGGCAAGATTGAGAAC	5022			
Db	5113	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCAGGCAAGATTGAGAAC	5172			
QY	5023	GGCTCTGCATCTGTGCGAGCCATGCCAGCAAGTGTCTTCTCGCTACACAGAAAAC	5082			
Db	5173	GGCTCTGCATCTGTGCGAGCCATGCCAGCAAGTGTCTTCTCGCTACACAGAAAAC	5232			
QY	5083	CTCAGAAAATACCTGCATCCGAAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC	5142			
Db	5233	CTCAGAAAATACCTGCATCCGAAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC	5292			
QY	5143	TTACCAATTTACAGTATCTCTCATTTGNAACCAATAAATTTACGAAATCGACATGAGCAG	5202			
Db	5293	TTACCAATTTACAGTATCTCTCATTTGNAACCAATAAATTTACGAAATCGACATGAGCAG	5352			
QY	5203	TACAGCTCCAGGAATTCCTGGATAGAAATGACCAATTCCTTGGCAGCTGTGTGTTGCC	5262			
Db	5353	TACAGCTCCAGGAATTCCTGGATAGAAATGACCAATTCCTTGGCAGCTGTGTGTTGCC	5412			
QY	5263	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACGCGAGGGCAGCAGAG	5322			
Db	5413	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACGCGAGGGCAGCAGAG	5472			
QY	5323	GAGTACTTGTGTGTTCACGAAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC	5382			
Db	5473	GAGTACTTGTGTGTTCACGAAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC	5532			
QY	5383	CGCAGACGATCTCAAGTGGAGTTCGCTTACCTTTTGGCCCTTTCCTTACAGAGAACCTTAT	5442			
Db	5533	CGCAGACGATCTCAAGTGGAGTTCGCTTACCTTTTGGCCCTTTCCTTACAGAGAACCTTAT	5592			
QY	5443	CTGTTTGTGACCACTTCAACTCACTCGAAGTAATTTAGATCCAGGCAAGCTCTCAGCA	5502			
Db	5593	CTGTTTGTGACCACTTCAACTCACTCGAAGTAATTTAGATCCAGGCAAGCTCTCAGCA	5652			
QY	5503	GGGACCTTGCCTGAGCGTACCTGGACATCCGAAACCCGGGCTACCTGGGCGCTTGCATT	5562			
Db	5653	GGGACCTTGCCTGAGCGTACCTGGACATCCGAAACCCGGGCTACCTGGGCGCTTGCATT	5712			
QY	5563	TCCTCAGAGCGATTTACTTTGGCGTCTCTCATACAGGATAAATTAAGGGTCAATTGTGTC	5622			
Db	5713	TCCTCAGAGCGATTTACTTTGGCGTCTCTCATACAGGATAAATTAAGGGTCAATTGTGTC	5772			
QY	5623	AAGGAAACCTCGTGAAGAGTCCGCGACCTGAAACACCAACCGGGGCGGCTCCACCTCCCGC	5682			

Db 778 CCAGATTATCATGGCTCTCTGAAGTGTGATGTGAACGGGATGGAAGGACCTAC 837
QY 841 GGCCTGGACTGTGACTGTGTGTCACTGGGCGTGTATGCTATGAGATGATTTATGGAGA 900
Db 838 GGCCTGGACTGTGACTGTGTGTCACTGGGCGTGTATGCTATGAGATGATTTATGGAGA 897
QY 901 TCCCCCTTCGACAGAGGAACTCTGCCGAACCTTCAATAACATTTATCAATTTCCAGCGG 960
Db 898 TCCCCCTTCGACAGAGGAACTCTGCCGAACCTTCAATAACATTTATCAATTTCCAGCGG 957
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGGAGTGTGAGTGTGATGATGATCAAAAGC 1020
Db 958 TTTTGAATTTCCAGATGACCCCAAGTGAAGGAGTGTGAGTGTGATGATGATCAAAAGC 1017
QY 1021 TTGTTGTGGGCGCAGAAAGAGAGATGTAAGTTTGAAGTCTTGTGCTGCATCTTCTTC 1080
Db 1018 TTGTTGTGGGCGCAGAAAGAGAGATGTAAGTTTGAAGTCTTGTGCTGCATCTTCTTC 1077
QY 1081 TCTAAAATTTGACTGGAACAAACATTTGTAAGTCTTGTGCTGCATCTTCTTC 1140
Db 1078 TCTAAAATTTGACTGGAACAAACATTTGTAAGTCTTGTGCTGCATCTTCTTC 1137
QY 1141 TCTGACGATGACACCTTCAATTTTGTGATGAACCCAGAGAGAAATTCGTGGTTTCATCTCT 1200
Db 1138 TCTGACGATGACACCTTCAATTTTGTGATGAACCCAGAGAGAAATTCGTGGTTTCATCTCT 1197
QY 1201 CGTGCACGCTGAGCCCTCAGCTCTCTCGGTGAAGAACTCCCGTTTGTGGGTTTTCG 1260
Db 1198 CGTGCACGCTGAGCCCTCAGCTCTCTCGGTGAAGAACTCCCGTTTGTGGGTTTTCG 1257
QY 1261 TACAGCAAGGCACTGGGGATTTCTGGTAGATCTGAGTCTGTTGTGTGGGTTCTGACTCC 1320
Db 1258 TACAGCAAGGCACTGGGGATTTCTGGTAGATCTGAGTCTGTTGTGTGGGTTCTGACTCC 1317
QY 1321 CTTGCCAAGACTAGCTCCATGGAAGAACTCTCTCAATAAGCAAGAGACTCAAGAC 1380
Db 1318 CTTGCCAAGACTAGCTCCATGGAAGAACTCTCTCAATAAGCAAGAGACTCAAGAC 1377
QY 1381 TCTCAGGCAAGTGTCAAGATGGAAGAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1440
Db 1378 TCTCAGGCAAGTGTCAAGATGGAAGAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1437
QY 1441 GAGGTGGAGGCTGTGCTTGTAGTGAAGAGGTTGAGCTGAAGCCCTCTGAGACTCAGAGA 1500
Db 1438 GAGGTGGAGGCTGTGCTTGTAGTGAAGAGGTTGAGCTGAAGCCCTCTGAGACTCAGAGA 1497
QY 1501 TCCCTCTGGAGCAGGACCTTGTCTACTTACATCACAGAAATGCAAGCTTAAAGCGAAGT 1560
Db 1498 TCCCTCTGGAGCAGGACCTTGTCTACTTACATCACAGAAATGCAAGCTTAAAGCGAAGT 1557
QY 1561 TTGGAGCAAGCAGGATGAGGTTGTCCTCAGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
Db 1558 TTGGAGCAAGCAGGATGAGGTTGTCCTCAGAGGATGACAAAGCACTGAGCTTCTCCAT 1617
QY 1621 GATATCAGAGAGCAGAGCCGAGCTCCAGAAATCAAGAGCAGAGTACAGGCTCAA 1680
Db 1618 GATATCAGAGAGCAGAGCCGAGCTCCAGAAATCAAGAGCAGAGTACAGGCTCAA 1677
QY 1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTTGAAGAGATCTTGTCTCAGCAAGAAGA 1740
Db 1678 GTGGAAGAAATGAGTTGATGATGAATCAGTTTGAAGAGATCTTGTCTCAGCAAGAAGA 1737
QY 1741 CGAGTGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTCTGAAGAAATTCAG 1800
Db 1738 CGAGTGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTCTGAAGAAATTCAG 1797
QY 1801 CGGAAGCGACAGAAATGCTCAGCAATAACTTGTGAAGCTTAAGGATCAAGGAGGCTGAA 1860
Db 1798 CGGAAGCGACAGAAATGCTCAGCAATAACTTGTGAAGCTTAAGGATCAGGGAAGCTGAA 1857
QY 1861 GTGGGAAATATGCGAAATGGAAGAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAG 1920

Db 1858 GTGGGAAATATGCGAAATGGAAGAGATCAATGTCTGACGACGAGCTCAAAATTCAGGAG 1917
QY 1921 CTCACAGAGAAAATCTGGAGAGGCT----- 1944
Db 1918 CTCACAGAGAAAATCTGGAGAGGCTGTAAAGCCAGCACGGAGGCCACCGAGCTGTGCAG 1977
QY 1945 -----GCAAGGAGCCGAGCGAGAGGAGCTGGAGAAAGCTGAGAAACCCGAGCTGTCTG 2052
Db 1978 AATATCGCCAGCAAAAGGAGCGAGCGAGAGGAGCTGGAGAAAGCTGTCAAGAACCGAGAG 2037
QY 1993 GATTCCTCTCAAGGCAATCAGAAAGAGCTGGTGAAGCTGAGGAAAGCCGCCATTTCTCTG 2052
Db 2038 GATTCCTCTCAAGGCAATCAGAAAGAGCTGGTGAAGCTGAGGAAAGCCGCCATTTCTCTG 2097
QY 2053 GAGAAACAAGGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2112
Db 2098 GAGAAACAAGGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2157
QY 2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATTAATTTCTGGAGCTCGAAGAG 2172
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATTAATTTCTGGAGCTCGAAGAG 2217
QY 2173 AAACATCGGAGGCGCCCAAGTCTCAGCCAGCAGCTTAGAAGTGCACCTGAAACAGAAAGAG 2232
Db 2218 AAACATCGGAGGCGCCCAAGTCTCAGCCAGCAGCTTAGAAGTGCACCTGAAACAGAAAGAG 2277
QY 2233 CAGCACTATGAGGAAAAGATTAAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC 2292
Db 2278 CAGCACTATGAGGAAAAGATTAAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC 2337
QY 2293 AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGGCCATGAGAAGGCAAA 2352
Db 2338 AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGGCCATGAGAAGGCAAA 2397
QY 2353 ATTCTCAGGCAACAGAAAGGCGATGATCAATGCTATGATTTCAAGATCAGATCCCTCGAA 2412
Db 2398 ATTCTCAGGCAACAGAAAGGCGATGATCAATGCTATGATTTCAAGATCAGATCCCTCGAA 2457
QY 2413 CAGAGATTTGGAACTGTCTGAGCCCAATAAATCTGCAGCAATAGCAGTCTTTTACC 2472
Db 2458 CAGAGATTTGGAACTGTCTGAGCCCAATAAATCTGCAGCAATAGCAGTCTTTTACC 2517
QY 2473 CAAAGGAACTGAAGGCGCCAAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTAC 2532
Db 2518 CAAAGGAACTGAAGGCGCCAAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTAC 2577
QY 2533 CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG 2592
Db 2578 CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG 2637
QY 2593 AAGATCAGCCACCAAGACCAAGTGAACAAGAAATCGGCTGTGGAACCTGGAGACAAAGATTG 2652
Db 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGAAATCGGCTGTGGAACCTGGAGACAAAGATTG 2697
QY 2653 CGGAGGCTCAGTCTAGAGACGAGGACGAGAAACTGGAGTCAAGCGCCAGCTCAGAGAG 2712
Db 2698 CGGAGGCTCAGTCTAGAGACGAGGACGAGAAACTGGAGTCAAGCGCCAGCTCAGAGAG 2757
QY 2713 CTAAGCTCTCTCTCAGAGGCGGAGTCAAGTTGACGCTGAGCTGAGCTGAGCTGAGCTGAG 2772
Db 2758 CTAAGCTCTCTCTCAGAGGCGGAGTCAAGTTGACGCTGAGCTGAGCTGAGCTGAGCTGAG 2817
QY 2773 GCCCTGAGAGCCAGCTTTGCCGCGGAAGACAGAGCTGGAAGAGACCAACAGCAGAAAGCT 2832
Db 2818 GCCCTGAGAGCCAGCTTTGCCGCGGAAGACAGAGCTGGAAGAGACCAACAGCAGAAAGCT 2877
QY 2833 GAAAGGAGATCCAGGCTCAGCGCACATGAGATGAATTCAGCCGCAAAATTTGATGCT 2892
Db 2878 GAAAGGAGATCCAGGCTCAGCGCACATGAGATGAATTCAGCCGCAAAATTTGATGCT 2937
QY 2893 CTTCTTAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAACACAGCTGACCCGAG 2952
Db 2938 CTTCTTAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAACACAGCTGACCCGAG 2997


```
Db 5113 CCCACATTTTGAAGCTCTCAAGGCGCCACTGTTTGGGCGAGCGAAGATTGAGAAC 5172
QY 5023 GGGCTCTGCATCTGTGACGCATGCCAGCAAGTGGTCAATCTCTCCGTACAAAGAAAC 5082
Db 5173 GGGCTCTGCATCTGTGACGCATGCCAGCAAGTGGTCAATCTCTCCGTACAAAGAAAC 5232
QY 5083 CTCAGCAATATCTGATCCGGAAGAGATAGAGACTCAGAGCCCTGACGTGTATCCAC 5142
Db 5233 CTCAGCAATATCTGATCCGGAAGAGATAGAGACTCAGAGCCCTGACGTGTATCCAC 5292
QY 5143 TTACCAATATCAGTATCTCTCAATTTGGAACCAATAAATTTAGCAATCGAATGAAGCAG 5202
Db 5293 TTACCAATATCAGTATCTCTCAATTTGGAACCAATAAATTTAGCAATCGAATGAAGCAG 5352
QY 5203 TACAGCTCGAGGAATCTCTGATAGATGATGACCAATCTCTGGCACCTGCTGTGTTGCC 5262
Db 5353 TACAGCTCGAGGAATCTCTGATAGATGATGACCAATCTCTGGCACCTGCTGTGTTGCC 5412
QY 5263 GGCCTCTTCCAAACAGCTTCCCTCTCTCAATCGTGCAGGTGAACAGCGCAGGGCAGCGAGAG 5322
Db 5413 GGCCTCTTCCAAACAGCTTCCCTCTCTCAATCGTGCAGGTGAACAGCGCAGGGCAGCGAGAG 5472
QY 5323 GAGTACTTCTGTTTCCAGAAATTTGAGAGTGTGCTGGATCTTACGGAAGACGTAGC 5382
Db 5473 GAGTACTTCTGTTTCCAGAAATTTGAGAGTGTGCTGGATCTTACGGAAGACGTAGC 5532
QY 5383 CGCACAGAGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGCCCTTACAGAAACCCCTAT 5442
Db 5533 CGCACAGAGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGCCCTTACAGAAACCCCTAT 5592
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGACGCTCTCAGCA 5502
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGACGCTCTCAGCA 5652
QY 5503 GGGACCCCTGCGCGAGCGTACCTGGACATCCCGAACCCGCGTACTCTGGCCCTGCCTAT 5562
Db 5653 GGGACCCCTGCGCGAGCGTACCTGGACATCCCGAACCCGCGTACTCTGGCCCTGCCTAT 5712
QY 5563 TCCTCAGGAGCGATTTACTTGGCGTCTCTCATACCAGGATAAATTAAGGGTCAATTTGCTGC 5622
Db 5713 TCCTCAGGAGCGATTTACTTGGCGTCTCTCATACCAGGATAAATTAAGGGTCAATTTGCTGC 5772
QY 5623 AAGGGAACCTTGTGAAGAGTCCGCGACTGAACACCCGCGCCCGTCCACCTCCCGC 5682
Db 5773 AAGGGAACCTTGTGAAGAGTCCGCGACTGAACACCCGCGCCCGTCCACCTCCCGC 5832
QY 5683 AGCAGCCCCCAACAGGAGGCCCAACCGTACACAGGACACATCACCAAGCGCGTGGCC 5742
Db 5833 AGCAGCCCCCAACAGGAGGCCCAACCGTACACAGGACACATCACCAAGCGCGTGGCC 5892
QY 5743 TCAGCCCAAGCGCCCGCAAGGCCCGCAGCCACCCCGGAGAGCAAGCAACCCCAACCGC 5802
Db 5893 TCAGCCCAAGCGCGCCCGCAAGGCCCGCAGCCACCCCGGAGAGCAAGCAACCCCAACCGC 5952
QY 5803 TACCGGAGGGGGAGCCGAGCTGCCAGGACAGTCTCTTGGCCGCCCGCCCTGGAGCGA 5862
Db 5953 TACCGGAGGGGGAGCCGAGCTGCCAGGACAGTCTCTTGGCCGCCCGCCCTGGAGCGA 6012
QY 5863 GAGAAGTCCCGCCGCGGATGCTTACAGCAGCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 5922
Db 6013 GAGAAGTCCCGCCGCGGATGCTTACAGCAGCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 6072
QY 5923 GAAGACAGGAGGGCGCGGCTGCTTGGGAGCGGTGAGGACCCCGCTGTGCCAGGTG 5982
Db 6073 GAAGACAGGAGGGCGCGGCTGCTTGGGAGCGGTGAGGACCCCGCTGTGCCAGGTG 6132
QY 5983 AACAAAGGGAAGAGGCA 5999
Db 6133 AACAAAGTGAAGGCA 6149
```

RESULT 11

```
US-10-028-946-3
; Sequence 3, Application US/10028946
; Publication No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5877
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-028-946-3
```

```
Query Match 87.0%; Score 5358; DB 14; Length 5877;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5634; Conservative 0; Mismatches 5; Indels 243; Gaps 3;
```

```
QY 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCCTTTGGATGCTGGTCTGTAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCCTTTGGATGCTGGTCTGTAACCCATT 60
QY 61 GCCAGCCGGGCTCCAGGCTGNAATCTCTTCCAGGGGAAACCCACCCTTTATGACTCAA 120
Db 61 GCCAGCCGGGCTCCAGGCTGNAATCTCTTCCAGGGGAAACCCACCCTTTATGACTCAA 120
QY 121 CAGCAGATGTCCTCTTTCCGAGAGAGGATATAGATGCCCTCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGTCCTCTTTCCGAGAGAGGATATAGATGCCCTCTTTGTTCTTTGAA 180
QY 181 GAATGCGCTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240
Db 181 GAATGCGCTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGCGCAAAGACTTCGAAAGTACA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGCGCAAAGACTTCGAAAGTACA 300
QY 301 AGTCTTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGTGTGTAAGAGAGAAAGCAACCGG 360
Db 301 AGTCTTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGTGTGTAAGAGAGAAAGCAACCGG 360
QY 361 GACATCTATGCTATGAAAGTGAAGAAAGGCTTTATTGGCCCGAGGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAAGTGAAGAAAGGCTTTATTGGCCCGAGGAGCAGGTTTCA 420
QY 421 TTTTGTGAGGAAGCGGGAACATATTATCTGAAAGCAAGCCCGTGTGATCCCAATT 480
Db 421 TTTTGTGAGGAAGCGGGAACATATTATCTGAAAGCAAGCCCGTGTGATCCCAATT 480
QY 481 CAGTATGCTTTTCAGGACAAAATCACTTTTATCTGATGAGGAATATCAGCTTGAGGG 540
Db 481 CAGTATGCTTTTCAGGACAAAATCACTTTTATCTGATGAGGAATATCAGCTTGAGGG 540
QY 541 GACTTGTCTCAGCTTTTGAATAGATATGAGGACCAAGTATGATAAAACCTGATCAGTTT 600
Db 541 GACTTGTCTCAGCTTTTGAATAGATATGAGGACCAAGTATGATAAAACCTGATCAGTTT 600
QY 601 TACCTAGCTGAGCTGATTTGGCTGTTTACAGCGTTCTATCTGATGGGATAGTGCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTGGCTGTTTACAGCGTTCTATCTGATGGGATAGTGCATCGA 660
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCGCACAGGACACATCAAGCTGGTGGATTT 720
Db 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCGCACAGGACACATCAAGCTGGTGGATTT 720
```

QY 721 GGATCTCCGCGAAATGAATTTAAAACAAGATGTTGAATCCAAATCCCGATTGGGACC 780
DB 721 GGATCTCCGCGAAATGAATTTAAAACAAGATGTTGAATCCAAATCCCGATTGGGACC 780
QY 781 CCAGATTACATGGCTCTGAAAGTCTGCTGATGAACCGGGATGGAAAGGACCTTAC 840
DB 781 CCAGATTACATGGCTCTGAAAGTCTGCTGATGAACCGGGATGGAAAGGACCTTAC 840
QY 841 GGCCTGAGCTGTGATCTGGTGGTCTAGTGGGCGTGATTCCTTATGAGATGATTTATGGGAGA 900
DB 841 GGCCTGAGCTGTGATCTGGTGGTCTAGTGGGCGTGATTCCTTATGAGATGATTTATGGGAGA 900
QY 901 TCCCCCTTCCAGAGGGAACCTTCGCGAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 960
DB 901 TCCCCCTTCCAGAGGGAACCTTCGCGAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 960
QY 961 TTTTGTGAATTTCCAGATGAGCCCAAGTGAAGTGAAGTCTTCTGATCTGATTTCAAAGC 1020
DB 961 TTTTGTGAATTTCCAGATGAGCCCAAGTGAAGTGAAGTCTTCTGATCTGATTTCAAAGC 1020
QY 1021 TTGTTGTGGCGGCAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCGCATCTTCTTCTTC 1080
DB 1021 TTGTTGTGGCGGCAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCGCATCTTCTTCTTC 1080
QY 1081 TCTAAAATTTGACTGGAACAACTTCGTAATCTCTCCGCCCTTCGTTCCACCCCTCAAG 1140
DB 1081 TCTAAAATTTGACTGGAACAACTTCGTAATCTCTCTCCGCCCTTCGTTCCACCCCTCAAG 1140
QY 1141 TCTGAAGTGAACCTTCAATTTTGAACCAAGAGAAATTCGTTGCTTCTTCTTCTTCT 1200
DB 1141 TCTGAAGTGAACCTTCAATTTTGAACCAAGAGAAATTCGTTGCTTCTTCTTCTTCT 1200
QY 1201 CCGTGCAGCTGAGCCCTCAGCTCTTCGCGTGAAGACTGCGCTTCTGCGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGCCCTCAGCTCTTCGCGTGAAGACTGCGCTTCTGCGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGGATTCCTGAGATCTGAGTCTGTTGTTGCTCGGCTCTGAGCTCC 1320
DB 1261 TACAGCAAGGCACTGGGGATTCCTGAGATCTGAGTCTGTTGTTGCTCGGCTCTGAGCTCC 1320
QY 1321 CCGTGCAGTACTGCTTCCATGGAAGAAACTTCTCATCAAGAGCAAGAGCTTACAGAC 1380
DB 1321 CCGTGCAGTACTGCTTCCATGGAAGAAACTTCTCATCAAGAGCAAGAGCTTACAGAC 1380
QY 1381 TCTCAGGACAAGTGTCAAGATGAGAGCAAGAAATCACCGGTTTACATCGGAGAGTGCA 1440
DB 1381 TCTCAGGACAAGTGTCAAGATGAGAGCAAGAAATCACCGGTTTACATCGGAGAGTGCA 1440
QY 1441 GAGTGGAGGCTGTGCTTGTAGTCAAGAGAGGAGTGAAGGCTGAAGGCTCTGAGACTCAGAGA 1500
DB 1441 GAGTGGAGGCTGTGCTTGTAGTCAAGAGAGGAGTGAAGGCTGAAGGCTCTGAGACTCAGAGA 1500
QY 1501 TCCCTCTCTGGAGCAGGACCTTCTACTCATCAAGATGAGTGTGCTTAAAGCGAAGT 1560
DB 1501 TCCCTCTCTGGAGCAGGACCTTCTACTCATCAAGATGAGTGTGCTTAAAGCGAAGT 1560
QY 1561 TTGGAGCAAGCAGGATGAGGTTGCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
DB 1561 TTGGAGCAAGCAGGATGAGGTTGCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
QY 1621 GATATCAGAGAGCAGAGCGGAGCTTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680
DB 1621 GATATCAGAGAGCAGAGCGGAGCTTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680
QY 1681 GTGGAGAAATGAGTTGATGATCAATGTTGGAAGGAGTCTTCTCTCAGCAAGAGA 1740
DB 1681 GTGGAGAAATGAGTTGATGATCAATGTTGGAAGGAGTCTTCTCTCAGCAAGAGA 1740
QY 1741 CCGAGTGAATCTTACGAATCTGAGCTGAGAGATCTCGGCTGCTGCTGAGAAATTTCAAG 1800
DB 1741 CCGAGTGAATCTTACGAATCTGAGCTGAGAGATCTCGGCTGCTGCTGAGAAATTTCAAG 1800
QY 1801 CGAAAGCGACAGAAATGTGAGCATAACTGTTGAGGCTTAAAGTCAAGGGAAGCCTGAA 1860

DB 1801 CGAAAGCGACAGAAATGTGAGCATAACTGTTGAGGCTTAAAGGCTCAAGGGAAGCCTGAA 1860
QY 1861 GTGGAGAGATATCGAAACTGGAGAGATCAATCTGAGCAGCAGCTCABAATTCAGGAG 1920
DB 1861 GTGGAGAGATATCGAAACTGGAGAGATCAATCTGAGCAGCAGCTCABAATTCAGGAG 1920
QY 1921 CTCCAGAGAAACTGGAGAGGCT----- 1944
DB 1921 CTCCAGAGAAACTGGAGAGGCTTAAAGCAGCAGCAGGAGCCACCGAGCTCTGCAG 1980
QY 1945 -----GCNAGGAGGAGCCGAGAGGAGCTGGNAGAGCTGCAGAACCGGAGAG 1992
DB 1981 AATATCCGCCAGGCAAGGAGCGAGAGGAGCTGGAGAAAGCTGCAGAACCGGAGAG 2040
QY 1993 GATTCTTCTGAGGAGCATCAGAAAGAGCTGGTGAAGCTGAGGAAACGCCGCCATTTCTCTG 2052
DB 2041 GATTCTTCTGAGGAGCATCAGAAAGAGCTGGTGAAGCTGAGGAAACGCCGCCATTTCTCTG 2100
QY 2053 GAGAAACAGGTTAAAGAGACTAGAGACCATGATGAGAGCTAGAGAAACAGACTGAAGGATGAC 2112
DB 2101 GAGAAACAGGTTAAAGAGACTAGAGACCATGATGAGAGCTAGAGAAACAGACTGAAGGATGAC 2160
QY 2113 ATCCAGACAAATCCCAACAGATCCAGAGATGCTGATTAATAATTTCTGAGCTCGAAGAG 2172
DB 2161 ATCCAGACAAATCCCAACAGATCCAGAGATGCTGATTAATAATTTCTGAGCTCGAAGAG 2220
QY 2173 AAAATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAGTGCACTTGAACAGAAAGAG 2232
DB 2221 AAAATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAGTGCACTTGAACAGAAAGAG 2280
QY 2233 CAGCACTATGAGGAAAGATTAAGTGTGAGCAATCAGATTAAGAGAGCTTGGCTGAC 2292
DB 2281 CAGCACTATGAGGAAAGATTAAGTGTGAGCAATCAGATTAAGAGAGCTTGGCTGAC 2340
QY 2293 AAGGAGACACTTGGAGAACATGATGACAGAGACACAGAGAGAGGCCCAATGAGAGAGGCAAA 2352
DB 2341 AAGGAGACACTTGGAGAACATGATGACAGAGACACAGAGAGAGGCCCAATGAGAGAGGCAAA 2400
QY 2353 ATTCTCAGCAGAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGGAA 2412
DB 2401 ATTCTCAGCAGAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGGAA 2460
QY 2413 CAGAGGATTTGGAACTGTCTGAAGCCAATAAATCTTGACCAATAGCAGCTCTTTTACC 2472
DB 2461 CAGAGGATTTGGAACTGTCTGAAGCCAATAAATCTTGACCAATAGCAGCTCTTTTACC 2520
QY 2473 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2532
DB 2521 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2580
QY 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAGAACCGAAAACTGGAGAGCAGCTGGAG 2592
DB 2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAGAACCGAAAACTGGAGAGCAGCTGGAG 2640
QY 2593 AAGATCAGCCACCAAGACCACTGACAGAAATCGGCTGCTGAACTGGAGACAAGATTG 2652
DB 2641 AAGATCAGCCACCAAGACCACTGACAGAAATCGGCTGCTGAACTGGAGACAAGATTG 2700
QY 2653 CGGAGGCTCAGTCTAGAGCAGGAGGAGGAGAACTGGAGCTCAAGCGCCAGCTCAGAGAG 2712
DB 2701 CGGAGGCTCAGTCTAGAGCAGGAGGAGGAGAACTGGAGCTCAAGCGCCAGCTCAGAGAG 2760
QY 2713 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGAGGCTGACCGGGG 2772
DB 2761 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGAGGCTGACCGGGG 2820
QY 2773 GCCTTGGAGAGCCAGCTTCCGAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGAGCT 2832
DB 2821 GCCTTGGAGAGCCAGCTTCCGAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGAGCT 2880
QY 2833 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2892

Db 2881 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940
QY 2893 CTTTCGTAAACAGCTGTACTGTATATCAGACCTTGGAGGACAGCTAAACAGCTGACCGAG 2952
Db 2941 CTTTCGTAAACAGCTGTACTGTATATCAGACCTTGGAGGACAGCTAAACAGCTGACCGAG 3000
QY 2953 GACACGCTGAACTCAACAAACCAAAATCTTCTACTTTGTCCAAAACAACCTGATGAGGCTTCT 3012
Db 3001 GACACGCTGAACTCAACAAACCAAAATCTTCTACTTTGTCCAAAACAACCTGATGAGGCTTCT 3060
QY 3013 GCGCCAAACGAGAGATTGTACAACTGCGAAGTGAAGTGACCACTCCGCGGAGATC 3072
Db 3061 GCGCCAAACGAGAGATTGTACAACTGCGAAGTGAAGTGACCACTCCGCGGAGATC 3120
QY 3073 ACGGAAACGAGATGACGCTTACCAAGCAGAGAAACGATGGAGGCTCTGAAGACCCAG 3132
Db 3121 ACGGAAACGAGATGACGCTTACCAAGCAGAGAAACGATGGAGGCTCTGAAGACCCAG 3180
QY 3133 TGCACCACTGCTGGAGAAACAGGTATCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3192
Db 3181 TGCACCACTGCTGGAGAAACAGGTATCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240
QY 3193 AAAGAGCGCAGTGGAGGAGCCTGGAGGAGCGTCTGGGTGATGAGAAATCCCAAGTTTGAG 3252
Db 3241 AAAGAGCGCAGTGGAGGAGCCTGGAGGAGCGTCTGGGTGATGAGAAATCCCAAGTTTGAG 3300
QY 3253 TGTTCGGGTTGAGAGCTGCAGAGATGCTTGGACACCGAGAAACAGAGCAGGCGGAGAGCC 3312
Db 3301 TGTTCGGGTTGAGAGCTGCAGAGATGCTTGGACACCGAGAAACAGAGCAGGCGGAGAGCC 3360
QY 3313 GATCAGCGGATCACCGAGTCTCGCAGGTGGTGGAGCTGCGAGTGAAGGAGCACAAAGGCT 3372
Db 3361 GATCAGCGGATCACCGAGTCTCGCAGGTGGTGGAGCTGCGAGTGAAGGAGCACAAAGGCT 3420
QY 3373 GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAACTGAAGGCCGAGAGCCTCTCT 3432
Db 3421 GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAACTGAAGGCCGAGAGCCTCTCT 3480
QY 3433 GACAACTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGNAATGCCCGAGCTTCA 3492
Db 3481 GACAACTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGNAATGCCCGAGCTTCA 3540
QY 3493 CAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAAGCAAGCCAAA 3552
Db 3541 CAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAAGCAAGCCAAA 3600
QY 3553 TTACAGCAGAGATGACCTGCAGAAAAATCAATTTTCGCTGACTCAAGGACTGCAA 3612
Db 3601 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGCTGACTCAAGGACTGCAA 3660
QY 3613 GAAGCTCTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGACCTTGGAGTATCAGCTG 3672
Db 3661 GAAGCTCTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGACCTTGGAGTATCAGCTG 3720
QY 3673 GAAAAATTTAGTCTTCTATTTCTATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAA 3732
Db 3721 GAAAAATTTAGTCTTCTATTTCTATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAA 3780
QY 3733 CAAACCAACTCATTAATTTCTGCAAGCCAAAATGGAACCACTGCTAAAAGAAAAG 3792
Db 3781 CAAACCAACTCATTAATTTCTGCAAGCCAAAATGGAACCACTGCTAAAAGAAAAG -- 3838
QY 3793 GGTTTATTTAGTCGACGGAAGAGGACCTGCTTTTACCCACACAGGTTCTCTGCAGTAC 3852
Db 3839 -----AGGTTCTCTGCAGTAC 3855
QY 3853 AATGAGCTGAGCTGGCCCTGGAGAGAGGAGAAAGCTCGTGTGCAAGCTAGAGGAGCC 3912
Db 3856 AATGAGCTGAGCTGGCCCTGGAGAGAGGAGAAAGCTCGTGTGCAAGCTAGAGGAGCC 3915
QY 3913 CTTTCAGAGACCGCATCGAGCTCCGCTCCGCCCCGGGAGGAGCTGCCACCGCAAGCA 3972
Db 3916 CTTTCAGAGACCGCATCGAGCTCCGCTCCGCCCCGGGAGGAGCTGCCACCGCAAGCA 3975

QY 3973 ACGGACCAACCCACACCCATCCAGCCAGCCACCCGAGGAGCAGATGCGCCATTCGCC 4032
Db 3976 ACGGACCAACCCACACCCATCCAGCCAGCCACCCGAGGAGCAGATGCGCCATTCGCC 4035
QY 4033 ATCGTGGGTGCGCAGAGCACCAGCCAGTGCATGAGCTGTGCGGCCCGCCATTCAGC 4092
Db 4036 ATCGTGGGTGCGCAGAGCACCAGCCAGTGCATGAGCTGTGCGGCCCGCCATTCAGC 4095
QY 4093 CCGAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTGGGGTCTTTAAGGAAACGATGCAC 4152
Db 4096 CCGAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTGGGGTCTTTAAGGAAACGATGCAC 4155
QY 4153 CACAATATTCCTACCGATTCACAGTAGACTGAACATGCGAGCCACAAGTGTGCTGTG 4212
Db 4156 CACAATATTCCTACCGATTCACAGTAGACTGAACATGCGAGCCACAAGTGTGCTGTG 4215
QY 4213 TGTCTGATACCGTGCACTTTGGACGCGCAGCATCCAAATGTCTCGAATGTGTCAGTGTATG 4272
Db 4216 TGTCTGATACCGTGCACTTTGGACGCGCAGCATCCAAATGTCTCGAATGTGTCAGTGTATG 4275
QY 4273 TGTCAACCCCAAGTGTCTCCAGTGTCTGCCAGCCACCTCGCGCTTGCCTGCTGAATATGCC 4332
Db 4276 TGTCAACCCCAAGTGTCTCCAGTGTCTGCCAGCCACCTCGCGCTTGCCTGCTGAATATGCC 4335
QY 4333 ACACACTTCCAGAGGCTTCTCCAGTGTCTGCCAGGTAACCTCCAGGCTCCAGACCAAG 4392
Db 4336 ACACACTTCCAGAGGCTTCTCCAGTGTCTGCCAGGTAACCTCCAGGCTCCAGACCAAG 4395
QY 4393 GAGCCACAGCAGCTTGCACCTGGAAGGGTGGATGAAGTGGCCAGGAAATAACAAACGA 4452
Db 4396 GAGCCACAGCAGCTTGCACCTGGAAGGGTGGATGAAGTGGCCAGGAAATAACAAACGA 4455
QY 4453 GGAAGCAAGCTGGGACAGGAAATCATTTCTCGAGGGATCAAAAGTCTCTATTAT 4512
Db 4456 GGAAGCAAGCTGGGACAGGAAATCATTTCTCGAGGGATCAAAAGTCTCTATTAT 4515
QY 4513 GACAAATGAAGCAGAGAGCTGGACAGAGCGCGTGGAAAGTTTGAAGTGTGCTGCCCTCCC 4572
Db 4516 GACAAATGAAGCAGAGAGCTGGACAGAGCGCGTGGAAAGTTTGAAGTGTGCTGCCCTCCC 4575
QY 4573 GACGGGATGATCTATTATTTGCTGCTTCCGAACTCGCAATACAGCCAAA 4632
Db 4576 GACGGGATGATCTATTATTTGCTGCTTCCGAACTCGCAATACAGCCAAA 4635
QY 4633 GCA----- 4635
Db 4636 GCAGATGTCOCATACATCTGAGATGGAATCTCACCCGCACACCACTGCTGCCCGGG 4695
QY 4636 ----- 4635
Db 4696 AGAACCTCTACTTGTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACCGCTTA 4755
QY 4636 -----GAAAGAGCAGAAAGCTGATGCTAAACTG 4662
Db 4756 GAATCAGTTGTGCGAGGTGGGAGAGTTTCTAGGGAAAAAGCAGAAAGCTGATGCTAAACTG 4815
QY 4663 CTTGGAATCTCCCTGCTGAAACTGAAAGTGAATGACCGTCTAGACATGAATGCAAGCTG 4722
Db 4816 CTTGGAATCTCCCTGCTGAAACTGAAAGTGAATGACCGTCTAGACATGAATGCAAGCTG 4875
QY 4723 CCCTTCAGTGAACAGGTGTTGTTGGTGGGCAACGAGGAGGGCTTACGCGCTGAATGTC 4782
Db 4876 CCCTTCAGTGAACAGGTGTTGTTGGTGGGCAACGAGGAGGGCTTACGCGCTGAATGTC 4935
QY 4783 TTGAAATCTCCCTAAACCATGCTCCAGGAAATTTGGAGCAGTCTTCCAAATTTATATATC 4842
Db 4936 TTGAAATCTCCCTAAACCATGCTCCAGGAAATTTGGAGCAGTCTTCCAAATTTATATATC 4995
QY 4843 AAGACCTTGGAGAGAGTACTCATATAGCAGGAGAGAGGGGCACTGTGTCTTGTGAC 4902
Db 4996 AAGACCTTGGAGAGAGTACTCATATAGCAGGAGAGAGGGGCACTGTGTCTTGTGAC 5055

```
QY 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4962
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115
QY 4963 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTTCGGGAGGCAAGATTGAAC 5022
Db 5116 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTTCGGGAGGCAAGATTGAAC 5175
QY 5023 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTCTCATTTCTCGCTACAAACGAAAC 5082
Db 5176 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTCTCATTTCTCGCTACAAACGAAAC 5235
QY 5083 CTCAGCAAAATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTGCGAGCTGTATCCAC 5142
Db 5236 CTCAGCAAAATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTGCGAGCTGTATCCAC 5295
QY 5143 TTCACCAATTTACAGTATCTCTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAG 5202
Db 5296 TTCACCAATTTACAGTATCTCTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAG 5355
QY 5203 TACAGCTCGAGGAATTCCTGGATAAGAAATGAACATTCCTTTGGCAGCTGCTGTGTTGCC 5262
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGAAATGAACATTCCTTTGGCAGCTGCTGTGTTGCC 5415
QY 5263 GCCTCTTCCACAGCTTCCCTGCTCAATCGTCAGTGAACCGCAGCGCAGCGAGAG 5322
Db 5416 GCCTCTTCCACAGCTTCCCTGCTCAATCGTCAGTGAACCGCAGCGCAGCGAGAG 5475
QY 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGAGTACG 5382
Db 5476 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGAGTACG 5535
QY 5383 CGACAGACATCTCAAGTGGAGTCTGCTACCTTTGGCCCTTACGAGAACCCCTAT 5442
Db 5536 CGACAGACATCTCAAGTGGAGTCTGCTACCTTTGGCCCTTACGAGAACCCCTAT 5595
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCTCAGCA 5502
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCTCAGCA 5555
QY 5503 GGAACCCCTCCCGAGGTAACCTGGACATCCCGAACCCCGGCTACCTGGGCCCTGCCATT 5562
Db 5656 GGAACCCCTCCCGAGGTAACCTGGACATCCCGAACCCCGGCTACCTGGGCCCTGCCATT 5715
QY 5563 TCCTCAGGAGCGATTACTTGGGCTCTCATACAGATAAATTAAAGGTCAATTGCTGC 5622
Db 5716 TCCTCAGGAGCGATTACTTGGGCTCTCATACAGATAAATTAAAGGTCAATTGCTGC 5775
QY 5623 AAGGGAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5682
Db 5776 AAGGGAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5835
QY 5683 AG 5684
Db 5836 AG 5837
```

RESULT 12

```
US-10-276-774-137
; Sequence 137, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1e. Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276, 774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
```

```
; SOFTWARE: Custom
; SEQ ID NO 137
; LENGTH: 3131
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-137

Query Match 41.6%; Score 2564.8; DB 13; Length 3131;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2781; Conservative 0; Mismatches 2; Indels 195; Gaps 2;

QY 2857 GCACATAGAGATGAATTCACGCGCAAAATTTGATGCTCTTGGTAACAGCTGTACTGTATC 2916
Db 1 GCACATAGAGATGAATTCACGCGCAAAATTTGATGCTCTTGGTAACAGCTGTACTGTATC 60
QY 2917 ACAGACCTGGAGGAGCAGCTAAACAGCTGACCGGAGCAACGCTGAACCTCAACCAACAA 2976
Db 61 ACAGACCTGGAGGAGCAGCTAAACAGCTGACCGGAGCAACGCTGAACCTCAACCAACAA 120
QY 2977 AACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGACGAGATTGTACAA 3036
Db 121 AACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGACGAGATTGTACAA 180
QY 3037 CTGCGAAGTGAAGTGGACCATCTCCGCGGAGATCACGGAACGAGAGATGCAGCTTACC 3096
Db 181 CTGCGAAGTGAAGTGGACCATCTCCGCGGAGATCACGGAACGAGAGATGCAGCTTACC 240
QY 3097 AGCCAGAACCAACGATGGAGGCTCTGAAGACCACTGACCGTCCAGGATCGAGGAGAACAGTC 3156
Db 241 AGCCAGAACCAACGATGGAGGCTCTGAAGACCACTGACCGTCCAGGATCGAGGAGAACAGTC 300
QY 3157 ATGGATTTGGAGGCCCTTAAACGATGCTGTAGAAAAGAGCGGAGTGGAGGCCCTGG 3216
Db 301 ATGGATTTGGAGGCCCTTAAACGATGCTGTAGAAAAGAGCGGAGTGGAGGCCCTGG 360
QY 3217 AGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGGAGTGTGGGTTTCGAGAGCTGCAGAGA 3276
Db 361 AGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGGAGTGTGGGTTTCGAGAGCTGCAGAGA 420
QY 3277 ATGTGGACACCGGAACACAGAGCGGCGAGCCGATCAGCGGATCAGCGAGTCTCGC 3336
Db 421 ATGTGGACACCGGAACACAGAGCGGCGAGCCGATCAGCGGATCAGCGAGTCTCGC 480
QY 3337 CAGTGTGGAGCTGGCAGTGAAGGACACAAGGCTCAGATTCTCGCTCTGCAGCAGGCT 3396
Db 481 CAGTGTGGAGCTGGCAGTGAAGGACACAAGGCTCAGATTCTCGCTCTGCAGCAGGCT 540
QY 3397 CTCAAAGACGAGAAGCTGAAGGCCGAGAGCCTCTCTGACAAGCTCAATGACCTGGAGAAG 3456
Db 541 CTCAAAGACGAGAAGCTGAAGGCCGAGAGCCTCTCTGACAAGCTCAATGACCTGGAGAAG 600
QY 3457 AAGCATGCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGGAACGA 3516
Db 601 AAGCATGCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGGAACGA 660
QY 3517 GAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAATTAACAGCAGCAGATGGACCTGCAG 3576
Db 661 GAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAATTAACAGCAGCAGATGGACCTGCAG 720
QY 3577 AAAAATCACATTTTCGCTCTGACTCAGGACTCAGAGACTTAGATCGGCTGATCTA 3636
Db 721 AAAAATCACATTTTCGCTCTGACTCAGGACTCAGAGACTTAGATCGGCTGATCTA 780
QY 3637 CTGAAGACAGAAAGAGTGAATGGAGTATCAGCTGGAAACCAATTCAGGTTCTCTATTCT 3696
Db 781 CTGAAGACAGAAAGAGTGAATGGAGTATCAGCTGGAAACCAATTCAGGTTCTCTATTCT 840
QY 3697 CATGAAAAGGTGAAAATGGAGGCACTATTCTCAACAAACCAAACTCATTGATTTCTG 3756
Db 841 CATGAAAAGGTGAAAATGGAGGCACTATTCTCAACAAACCAAACTCATTGATTTCTG 900
QY 3757 CAAGCCAAAATGGACCAACCTGCTTAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAG 3816
Db 3757 CAAGCCAAAATGGACCAACCTGCTTAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAG 3816
```



```

; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Eisinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ott, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol B. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Barghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-482C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 7
; LENGTH: 2542
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(2542)
; US-10-262-511-7

Query Match 39.2%; Score 2414; DB 13; Length 2542;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2475; Conservative 0; Mismatches 5; Indels 48; Gaps 1;

QY 1846 CAAGGGAAGCCTGAAGTGGAGAGATATGCGAACTGGAGAGATCAATGCTGAGCAGCAG 1905
Db 11 CAGGGGAGCCTGAAGTGGAGAGATATGCGAACTGGAGAGATCAATGCTGAGCAGCAG 70
QY 1906 CTCAAAATTCAGAGCTCCAGAGAACTGGAGAGCCT----- 1944

```



```
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 5
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(2497)
; US-10-262-511-5

Query Match      37.6%; Score 2317.2; DB 13; Length 2497;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 2432; Conservative 0; Mismatches 3; Indels 93; Gaps 2;

QY 1846 CAAGGGAAGCTGAAGTGGAGATATGCGAAACTGGAGAGATCAATCTGAGCAGCAG 1905
Db 11 CAGGGGAAGCTGAAGTGGAGATATGCGAAACTGGAGAGATCAATCTGAGCAGCAG 70
QY 1906 CTCAAAATTCAGAGCTCCAAAGAAACTGGAGAGGCT----- 1944
Db 71 CTCAAAATTCAGAGCTCCAAAGAAACTGGAGAGGCTGTAAAGCCAGCGAGGCC 130
QY 1945 -----GAAAGGAGCGAGCCGAGAGGGAGCTGGAGAG 1977
Db 131 ACGAGCTGCTGAGAATATCCCGAGGCAAAAGAGCGAGCGAGAGGAGCTGGAGAG 190
QY 1978 CTCAGAAACGAGAGGATCTCTGAGGCGCATCAGAAAGAGCTGTGGAAGCTGAGGA 2037
Db 191 CTCAGAAACGAGAGGATCTCTGAGGCGCATCAGAAAGAGCTGTGGAAGCTGAGGA 250
QY 2038 CGCGCGCATCTCTGAGAGAAAGAGTAAAGAGACTAGAGACCATGAGGCTAGAGAAAC 2097
Db 251 CGCGCGCATCTCTGAGAGAAAGAGTAAAGAGACTAGAGACCATGAGGCTAGAGAAAC 310
QY 2098 AGACTGAGGATGACATCCAGACAAATCCCAACGATCCAGCAGTGGCTGATTAATTT 2157
Db 311 AGACTGAGGATGACATCCAGACAAATCCCAACGATCCAGCAGTGGCTGATTAATTT 370
QY 2158 CTGAGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCAGCACCATAGAAAGTGCAC 2217
Db 371 CTGAGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCAGCACCATAGAAAGTGCAC 430
QY 2218 CTGAACAGAAAGAGCAGCAGCTATGAGAAAGATTAAGTGTGGACAAATCAGATAAG 2277
Db 431 CTGAACAGAAAGAGCAGCAGCTATGAGAAAGATTAAGTGTGGACAAATCAGATAAG 490
QY 2278 AAAGACTGCTGCTCACAAGAGAGACACTGGAGAAACATGATCAGAGACACGAGAGGAGGCC 2337
Db 491 AAAGACTGCTGCTCACAAGAGAGACACTGGAGAAACATGATGAGAGACACGAGAGGAGGCC 550
QY 2338 CATGAGAAAGGCGAAAATCTCAGCGAAACAGAGGCGATCAATGCTATGATTCCTCAAG 2397
Db 551 CATGAGAAAGGCGAAAATCTCAGCGAAACAGAGGCGATCAATGCTATGATTCCTCAAG 610
QY 2398 ATCAGATCCCTGGAACAGAGGATTTGGAACTGCTGAGAGCAATTAATCTCAGCAAT 2457
Db 611 ATCAGATCCCTGGAACAGAGGATTTGGAACTGCTGAGAGCAATTAATCTCAGCAAT 670
QY 2458 AGCAGTCTTTTACCAGAGGACATGAAGGCCCAAGAGAGATGATTTCTGAATCAGG 2517
Db 671 AGCAGTCTTTTACCAGAGGACATGAAGGCCCAAGAGAGATGATTTCTGAATCAGG 730
QY 2518 CAACAGAAATTTTACCTGGAGACACAGGCTGGAGTTGGAGGCCCGAGAACCGAAACTG 2577
Db 731 CAACAGAAATTTTACCTGGAGACACAGGCTGGAGTTGGAGGCCCGAGAACCGAAACTG 790

2578 GAGGAGAGCTGGAGAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTCTCGAA 2637
791 GAGGAGAGCTGGAGAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTCTCGAA 850
2638 CTGAGAGCAAGATTGCGGGAGGTGAGTCTAGAGACCAAGAGAGAGAGAACTGGAGCTCAAG 2697
851 CTGAGAGCAAGATTGCGGGAGGTGAGTCTAGAGACCAAGAGAGAGAGAACTGGAGCTCAAG 910
2698 GCGCAGCTCAAGAGCTGACAGCTCTCCCTGAGAGAGCGGAGTCAAGCTTGAAGCCCTG 2757
911 GCGCAGCTCAAGAGCTGACAGCTCTCCCTGAGAGAGCGGAGTCAAGCTTGAAGCCCTG 970
2758 CAGGCTGACCGGCGGCCCTGAGAGAGCGGCTTCCGAGGCGGAGAGAGAGAGAGAG 2817
971 CAGGCTGACCGGCGGCCCTGAGAGAGCGGCTTCCGAGGCGGAGAGAGAGAGAGAG 1030
2818 ACCAGAGAGAGCTGAAGAGAGATCCAGGCACTCAGGCACTACAGACATAGAGATGAATCCAG 2877
1031 ACCAGAGAGAGCTGAAGAGAGATCCAGGCACTCAGGCACTACAGACATAGAGATGAATCCAG 1090
2878 CGCAAAATTTGATGCTCTCTGTAAGAGCTGTACTGTAACTCAGACCTGGAGGAGCAGCTA 2937
1091 CGCAAAATTTGATGCTCTCTGTAAGAGCTGTACTGTAACTCAGACCTGGAGGAGCAGCTA 1150
2938 AACGAGCTGACCGGAGGACAAAGCTGAATCAACCAACCAAACTTCTACTTTGTCCAAACAA 2997
1151 AACGAGCTGACCGGAGGACAAAGCTGAATCAACCAACCAAACTTCTACTTTGTCCAAACAA 1210
2998 CTCGATGAGGCTTCTGCGGCAACGAGAGATGTGTAACCTGCGAAGTGAAGTGGACAT 3057
1211 CTCGATGAGGCTTCTGCGGCAACGAGAGATGTGTAACCTGCGAAGTGAAGTGGACAT 1270
3058 CTCGCGCGGAGATCAAGGAAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGGAG 3117
1271 CTCGCGCGGAGATCAAGGAAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGGAG 1330
3118 GCTCTGAAGACCAAGTGCACATGCTGGAGAAACAGGTCAATGATTTGGAGGCCCTTAAC 3177
1331 GCTCTGAAGACCAAGTGCACATGCTGGAGAAACAGGTCAATGATTTGGAGGCCCTTAAC 1390
3178 GATGAGCTGCTAGAAAGAGCGGAGCTGGAGGCGCTGGAGGAGCGCTCTGGGTGATGAG 3237
1391 GATGAGCTGCTAGAAAGAGCGGAGCTGGAGGCGCTGGAGGAGCGCTCTGGGTGATGAG 1450
3238 AAATCCCAGTTTGAAGTGTGCGGTTGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAG 3297
1451 AAATCCCAGTTTGAAGTGTGCGGTTGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAG 1510
3298 AGCAGGCGGAGCGGATCAGCGGATCAGCGGATCAGCGGATGCTGGAGGCTGGAGCTGGAGTG 3357
1511 AGCAGGCGGAGCGGATCAGCGGATCAGCGGATCAGCGGATGCTGGAGGCTGGAGCTGGAGTG 1570
3358 AAGGAGCACAAGGCTGAGATTCTCGCTCTCAGCAGGCTCTCAAGAGCAGAGAGCTGAAG 3417
1571 AAGGAGCACAAGGCTGAGATTCTCGCTCTCAGCAGGCTCTCAAGAGCAGAGAGCTGAAG 1630
3418 GCGAGAGCCTCTCTGACAGAGCTCAATGAGCTGGAGAGAGAGATGCTATGCTGAATG 3477
1631 GCGAGAGCCTCTCTGACAGAGCTCAATGAGCTGGAGAGAGAGATGCTATGCTGAATG 1690
3478 AATGCCGAGAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTG 3537
1691 AATGCCGAGAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTG 1750
3538 GAAGAGCAAGCCAAATTTACAGCAGCAGATGGAAGCTGGAGAGAGAGATCAATTTTCCGCTG 3597
1751 GAAGAGCAAGCCAAATTTACAGCAGCAGATGGAAGCTGGAGAGAGAGATCAATTTTCCGCTG 1810
3598 ACTCAAGGAGCTGCAAGAGGCTCTAGATCGGGCTGATCTACTGAGAGCAGAGAGAGAGTGAC 3657
1811 ACTCAAGGAGCTGCAAGAGGCTCTAGATCGGGCTGATCTACTGAGAGCAGAGAGAGAGTGAC 1870
```

```
QY 3658 TTGGAGTATCAGTGGAAAAACATTGAGGTTCTCTATTCTCATGAAAAGGTGAAATGGAA 3717
Db 1871 TTGGAGTATCAGTGGAAAAACATTGAGGTTCTCTATTCTCATGAAAAGGTGAAATGGAA 1930
QY 3718 GGCATATTCTCAACAACCAAACTCATGTTTCTTCTGCAAGCCAAATGACCAAGCT 3777
Db 1931 GGCATATTCTCAACAACCAAACTCATGTTTCTTCTGCAAGCCAAATGACCAAGCT 1990
QY 3778 GCTAAAAAGAAAAGGGTTTATTATGTCGACGGAAGAGGACCTTCTTTTACCCACACAG 3837
Db 1991 GCTAAAAAGAAA-----AG 2005
QY 3838 GTTCTCTGCACTCAATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCA 3897
Db 2006 GTTCTCTGCACTCAATGAGTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCA 2065
QY 3898 GAGCTAGAGGAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCGCCCGGGAGGAAGCT 3957
Db 2066 GAGCTAGAGGAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCGCCCGGGAGGAAGCT 2125
QY 3958 GCCACCCGCAAGCAACGAGCACCCACACCCATCCAGCCAGCCACCGGAGGAGGAGCAG 4017
Db 2126 GCCACCCGCAAGCAACGAGCACCCACACCCATCCAGCCAGCCACCGGAGGAGGAGCAG 2185
QY 4018 ATCGCATGTCCGCTCGTGGTCCGAGAGCAGCAGCCAGTGCATGAGCTGCTG 4077
Db 2186 ATCGCATGTCCGCTCGTGGTCCGAGAGCAGCAGCCAGTGCATGAGCTGCTG 2245
QY 4078 GCCCGGCCATCCAGCCGCAAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 4137
Db 2246 GCCCGGCCATCCAGCCGCAAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 2305
QY 4138 AAGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4197
Db 2306 AAGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2365
QY 4198 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4257
Db 2366 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2425
QY 4258 GAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4317
Db 2426 GAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2485
QY 4318 CCTGCTGA 4325
Db 2486 CCTGCTGA 2493
```

RESULT 15

```
US-10-262-511-3
; Sequence 3, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
```

```
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1870)
US-10-262-511-3
```

Query Match 29.9%; Score 1841; DB 13; Length 1870;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 ATGTTGAAGTTCAAATATGGAGCGCGAATCCCTTTGGATGCTGCTGCTGAACCCATT 60
Db 14 ATGTTGAAGTTCAAATATGGAGCGCGAATCCCTTTGGATGCTGCTGCTGAACCCATT 73
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
Db 74 GCCAGCCGGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 133
QY 121 CAGCAGATGTCCTCTTTCCGAGAGGATATAGATCCCTCTTTGTTCTTTTGA 180
Db 134 CAGCAGATGTCCTCTTTCCGAGAGGATATAGATCCCTCTTTGTTCTTTGAA 193
QY 181 GAATGCACTAGCTGCTCTGATGAAGATTAAAGCAGTGAACACTTTGCCGAAGTAT 240
Db 194 GAATGCACTAGCTGCTCTGATGAAGATTAAAGCAGTGAACACTTTGCCGAAGTAT 253
QY 241 TCCGACACCATAGCTGATGATTACAGGAGCTCCAGCTTCGGCAAGGACTTCGAGTCA 300
Db 254 TCCGACACCATAGCTGATGATTACAGGAGCTCCAGCTTCGGCAAGGACTTCGAGTCA 313
QY 301 AGTCTTGTAGGTTGTGGTCTACTTTGCTGAAGTGCAGTGGTAAAGAGAGAAAGCAACCGGG 360
Db 314 AGTCTTGTAGGTTGTGGTCTACTTTGCTGAAGTGCAGTGGTAAAGAGAGAAAGCAACCGGG 373
```

QY	361	GACATCTATGCTATGAAGAGTGATGAAGAAGAGGCTTTATTGGCCCAAGGACAGGTTTCA	420
DB	374	GACATCTATGCTATGAAGAGTGATGAAGAAGAGGCTTTATTGGCCCAAGGACAGGTTTCA	433
QY	421	TTTTTTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATT	480
DB	434	TTTTTTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATT	493
QY	481	CAGTATGCTTTTCAGGACAAAAATCACTTTATCTGATGGAGGAATATCAGCCTGGAGGG	540
DB	494	CAGTATGCTTTTCAGGACAAAAATCACTTTATCTGATGGAGGAATATCAGCCTGGAGGG	553
QY	541	GACTTGGCTGCACATTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTTGATACAGTTT	600
DB	554	GACTTGGCTGCACATTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTTGATACAGTTT	613
QY	601	TACCTAGCTGAGCTGATTTTGGCTGTTTCAACGCTTTCATCTGATGGGATAGTGTCATCGA	660
DB	614	TACCTAGCTGAGCTGATTTTGGCTGTTTCAACGCTTTCATCTGATGGGATAGTGTCATCGA	673
QY	661	GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGACACATCAAGCTGGTGGATTTT	720
DB	674	GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGACACATCAAGCTGGTGGATTTT	733
QY	721	GGATCTGCCGCGAAAAATGAATTTCAAAACAAGATGGTGAATGCCAACTCCCGATTGGGACC	780
DB	734	GGATCTGCCGCGAAAAATGAATTTCAAAACAAGATGGTGAATGCCAACTCCCGATTGGGACC	793
QY	781	CCAGATTACATGCTCTCGTAAGTGCTGACTGTGATGAACGGGATGAAAGGCACCTAC	840
DB	794	CCAGATTACATGCTCTCGTAAGTGCTGACTGTGATGAACGGGATGAAAGGCACCTAC	853
QY	841	GGCTTGCACTGTGACTGTGTGTCAGTGGGCGTGATTCCTATGAGATGATTTATGGGAGA	900
DB	854	GGCCTGGACTGTGACTGTGTGTCAGTGGGCGTGATTCCTATGAGATGATTTATGGGAGA	913
QY	901	TCCGCCCTTCGACAGAGGNAACCTCTGCCAGAACCTTCAATAACATTTAGATTTCCAGCGG	960
DB	914	TCCGCCCTTCGACAGAGGNAACCTCTGCCAGAACCTTCAATAACATTTAGATTTCCAGCGG	973
QY	961	TTTTTTGAAATTTCCAGATGACCCCAAAGTGAGCAGTGACTTCTTGATCTGATTCAAAGC	1020
DB	974	TTTTTTGAAATTTCCAGATGACCCCAAAGTGAGCAGTGACTTCTTGATCTGATTCAAAGC	1033
QY	1021	TTGTTGTGGCCGACAGAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATTCCTTTCTTC	1080
DB	1034	TTGTTGTGGCCGACAGAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATTCCTTTCTTC	1093
QY	1081	TCTAAAATTTGACTGGAAACAATTCGTAACTCTCCTCCCCCTTCGTTCCACCCCTCAAG	1140
DB	1094	TCTAAAATTTGACTGGAAACAATTCGTAACTCTCCTCCCCCTTCGTTCCACCCCTCAAG	1153
QY	1141	TCTGACGATGACACTCAATTTTGTATGAAACCGAGAGAATTCGTGGGTTTCACTCTCT	1200
DB	1154	TCTGACGATGACACTCAATTTTGTATGAAACCGAGAGAATTCGTGGGTTTCACTCTCT	1213
QY	1201	CCGTGCCAGCTGAGCCCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGGTTTTCG	1260
DB	1214	CCGTGCCAGCTGAGCCCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGGTTTTCG	1273
QY	1261	TACAGCAAGGCACTGGGGATTTCTTGTGTAGATCTGAGTCTGTTGTGCGGTTCTGGACTCC	1320
DB	1274	TACAGCAAGGCACTGGGGATTTCTTGTGTAGATCTGAGTCTGTTGTGCGGTTCTGGACTCC	1333
QY	1321	CCTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1380
DB	1334	CCTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC	1393
QY	1381	TCTCAGGACAAGTGTCAAGAATGGAGCAGGAATGACCCGGTTACATCGGAGGTGTCA	1440
DB	1394	TCTCAGGACAAGTGTCAAGAATGGAGCAGGAATGACCCGGTTACATCGGAGGTGTCA	1453

Search completed: July 3, 2004, 23:37:58
Job time : 1770 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 12:14:48 ; Search time 279 Seconds
(without alignments)
12250.690 Million cell updates/sec

Title: US-10-017-216-3
Perfect score: 6159
Sequence: 1 atgttgagttccaatatgg.....ttctgagacagattattgc 6159

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A-COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B-COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A-COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B-COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCTUS-COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1391.8	22.6	1515	4	US-09-804-471A-1
2	1391.8	22.6	1515	4	US-10-238-709-1
3	750	12.2	1133	4	US-09-916-204-1
4	258	4.2	258	4	US-09-016-434-513
5	220.8	3.6	2726	2	US-08-422-699A-12
6	220.8	3.6	2726	2	US-08-422-706B-12
7	214	3.5	2511	2	US-08-422-699A-8
8	214	3.5	2511	2	US-08-422-706B-8
9	213.6	3.5	3182	1	US-08-484-044-11
10	205	3.3	174493	4	US-09-804-471A-3
11	205	3.3	174493	4	US-10-238-709-3
12	189.2	3.1	2706	2	US-08-630-822A-61
13	189.2	3.1	2706	2	US-09-005-069-61
14	189.2	3.1	2706	4	US-09-171-156A-20
15	189.2	3.1	2706	4	US-09-004-730A-20
16	189.2	3.1	2706	4	US-08-981-799A-20
17	181.2	2.9	4363	2	US-08-685-576-5
18	179.6	2.9	4848	4	US-09-976-594-295
19	178.8	2.9	48763	4	US-09-916-204-3
20	172.2	2.8	4055	4	US-09-016-434-1105
21	172.2	2.8	4739	3	US-08-685-871-1
22	168.4	2.7	5033	2	US-08-685-576-2
23	150.2	2.4	3323	2	US-08-422-699A-10
24	150.2	2.4	3323	2	US-08-422-706B-10
25	122.6	2.0	3155	4	US-09-442-100-7
26	122.6	2.0	3155	4	US-08-939-106-7
27	122.6	2.0	3155	4	US-09-442-102-7

28	122.2	2.0	1498	4	US-09-509-902A-6	Sequence 6, Appli
29	122.2	2.0	1961	4	US-09-509-902A-15	Sequence 15, Appli
30	122.2	2.0	5276	4	US-09-233-857-2	Sequence 2, Appli
31	117.8	1.9	3018	2	US-08-860-150-6	Sequence 6, Appli
32	117.8	1.9	3018	3	US-09-338-132-6	Sequence 6, Appli
33	111.4	1.8	1935	3	US-08-878-989-11	Sequence 11, Appli
34	111.4	1.8	1935	3	US-09-272-796-11	Sequence 11, Appli
35	109.8	1.8	3213	4	US-09-442-100-5	Sequence 5, Appli
36	109.8	1.8	3213	4	US-08-939-106-5	Sequence 5, Appli
37	109.8	1.8	3213	4	US-09-442-102-5	Sequence 5, Appli
38	105	1.7	3489	2	US-08-728-323A-1	Sequence 1, Appli
39	105	1.7	3489	4	US-09-298-568-1	Sequence 1, Appli
40	105	1.7	3489	4	US-09-410-399-1	Sequence 1, Appli
C 41	105	1.7	32207	2	US-08-770-379-20	Sequence 20, Appli
C 42	105	1.7	32207	3	US-08-757-669A-20	Sequence 20, Appli
C 43	105	1.7	32207	4	US-09-230-371A-20	Sequence 20, Appli
44	104.8	1.7	2160	3	US-09-588-256-1	Sequence 1, Appli
45	101.4	1.6	5720	4	US-09-442-100-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-804-471A-1
; Sequence 1, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Human
US-09-804-471A-1

Query Match		22.6%	Score 1391.8;	DB 4;	Length 1515;
Best Local Similarity		99.5%	Pred. No. 0;		
Matches 1396;		Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
QY	1	ATGTTGAAGTTCAATATGGAGCGCGGAATCTTTGGATGCTGCTGCTGCTGAACCCATT	60		
DB	7	ATGTTGAAGTTCAATATGGAGCGCGGAATCTTTGGATGCTGCTGCTGCTGAACCCATT	66		
QY	61	GCCAGCGGGCCTCCAGGCTGAATCTTCTTCCAGGGGAAACACCCCTTTATGACTCAA	120		
DB	57	GCCAGCGGGCCTCCAGGCTGAATCTTCTTCCAGGGGAAACACCCCTTTATGACTCAA	126		
QY	121	CAGCAGATGCTCTCTTTCCGAGAGGAGGATATAGATGCCCTCTTTGCTCTTTGAA	180		
DB	127	CAGCAGATGCTCTCTTTCCGAGAGGAGGATATAGATGCCCTCTTTGCTCTTTGAA	186		
QY	181	GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCACTTTGTCGGGAAGTAT	240		
DB	187	GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCACTTTGTCGGGAAGTAT	246		
QY	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTCAGA	300		
DB	247	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTCAGA	306		
QY	301	AGTCTTGTAGTTTGTGCTCACTTGTGAGTCAGCTGCTTAAGAGAGAGAACCGGG	360		
DB	307	AGTCTTGTAGTTTGTGCTCACTTGTGAGTCAGCTGCTTAAGAGAGAGAACCGGG	366		
QY	361	GACATCATGTATGAAAGTATGAAGAAGAGGCTTTATTGGCCCGAGGAGGAGTTTCA	420		

APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 513:
SEQUENCE CHARACTERISTICS:
LENGTH: 258 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINON01
CLONE: 2290031
US-09-016-434-513

Query Match 4.2%; Score 258; DB 4; Length 258;

Best Local Similarity 100.0%; Pred. No. 3.4e-60;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5051 GCAAAGTCGTCAATTCCTCCGCTACACGAAACCTCAGCAAAATCTGCTCCGGAAGAGA 5110
DB 1 GCAAAGTCGTCAATTCCTCCGCTACACGAAACCTCAGCAAAATCTGCTCCGGAAGAGA 60
QY 5111 TAGAGACCTCAGAGCCCTGAGCTGTATCCACTTACCAATTAAGTATCCTCATTTGAA 5170
DB 61 TAGAGACCTCAGAGCCCTGAGCTGTATCCACTTACCAATTAAGTATCCTCATTTGAA 120
QY 5171 CCAATAAATCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTCCTGGATAAGA 5230
DB 121 CCAATAAATCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTCCTGGATAAGA 180
QY 5231 ATGACCATTCCTTGGACCTGCTGTGTTTCCGCGCTCTTCCAAAGCTTCCCTGTCTCAA 5290
DB 181 ATGACCATTCCTTGGACCTGCTGTGTTTCCGCGCTCTTCCAAAGCTTCCCTGTCTCAA 240
QY 5291 TCGTCAGGTGACACGG 5308
DB 241 TCGTCAGGTGACACGG 258

RESULT 5

US-08-422-699A-12
Sequence 12, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-422-699A-12

Query Match 3.6%; Score 220.8; DB 2; Length 2726;

Best Local Similarity 53.7%; Pred. No. 2e-49;

Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;

QY 194 CTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTGTCCGGAAGTATTCGACACCATAG 253
DB 243 CCGAAGTGGCCAGGACAGTAGCTGGCCGACTTCTTGCACTGGCGGAGCCCATGTCG 302
QY 254 CTGAGTTACAGGAGCTCCAGCCTTCGGCAAGAGCACTTCGAAGTCAGAAAGTCTTTGAGGTT 313
DB 303 TGAGGCTTAAGGAGGTCCGACTCGAGAGGACGACTTCGAGATTCTGAAGGTGATCGGAC 362
QY 314 GTGTCACATTTGCTGAAGTGCAGGTGTAAGAGAGAAAGCAACCGGGGACATCTATGCTA 373
DB 363 CGCGGGCTTCACGAGGAGTAGCGGTAGTAGAGATGAAGCAGCGGCGGAGGTCATGCCA 422
QY 374 TGAAAGTGAATGAAGAAAGGCTTTATTTGGCCAGAGCAGGTTTCAATTTTGAAGAAG 433
DB 423 TGAAGATCATGAACAAGTGGACATCTGTAAGAGGGCGGAGGTGTCGTGCTTCCGTGAGG 482
QY 434 AGCGGAACATATTATCTCGNAGCACAAGCCCGTGGATCCCAATTAAGTATGCTTTC 493
DB 483 AGAGGACGTTGTTGGTGAATGGGACCGGGGTGGATCAAGCAGTGCACCTTCGCTTCC 542
QY 494 AGGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTCGAGGGGAGCTTGTGTCTAC 553
DB 543 AGGATGAGAACTACCTGTACCTGTCATGGAGTATTACGTCGGCGGGGACCTGCTGACAC 602
QY 554 TTTTGAATAGATAGAGACCAAGTATGATAAACCCTGATACAGTTTACCTAGCTGAGC 613
DB 603 TGCTGAGCAAGTTTGGGAGCGGATTCGCGCCGAGATGCGCGCTTCTACCTGGGGAGA 662
QY 614 TGATTTTGGCTGTTTCAAGCGGTTCACTCTGATGGGATACGTGATCGAGACATCAAGCCTG 673

Db 663 TTGTCATGGCCATAGACTCGGTGCACCGGCTTGGCTACGTGCACAGGACATCAAAACCG 722
Qy 674 AGAACATCTCTGTTGACCGCACAGGACACATCAAGCTGTGGATTTTGGATCTGCGCGA 733
Db 723 ACAACATCTCTGTTGACCGCACAGGACACATCAAGCTGTGGATTTTGGATCTGCGCTCA 782
Qy 734 AAATGAATTCACAAAGATGGTGAATGCCAAACTCCCGATTGGGACCCAGATTTACATGG 793
Db 783 AGCTGCGGCGAGATGGAACGGTGGCTGGCTGTGGGACACCCAGACTACCTGT 842
Qy 794 CTCCTGAAGTCTGA----CTGTGATGAACGGGATGGAAGGACCTACGCGCTGGACT 850
Db 843 CCCCCGAGATCTCTGACGCTGTGGGCGGTGGGCTGGGACAGGACCTACGCGCGCGAGT 902
Qy 851 GTGACTGGTGTGATGCGGCGGTGATTCCTATCAGATGATTTATGGAGATCCCTCTCG 910
Db 903 GTGACTGGTGGCGCTGGTGTGATTCCTATCAGATGATTTATGGAGATCCCTCTCG 962
Qy 911 CAGAGGAACTCTGCGCAGAACCTTCAATAACATTAATGAATTCAGCGGTTTTGAAAT 970
Db 963 ACGGGATTCACGGCGGAGACCTATGGCAAGATCGTCCACTACAGGAGACCTCTCTC 1022
Qy 971 TTCCAGATGACCCCAAAAGTGAAGT--GACITTTCTGTATCTGATTCAGATTCAGGCTTGT 1027
Db 1023 TGCGCTGGTGGAGGAGGAGGCTCTGAGGAGGCTCGAGACTTCATTCAGCGTGTCTGT 1082
Qy 1028 CGGCGCAGAGAGAGACTGAAATTTGAAGTCTTTGCTGCT-----CATCTTTCT 1078
Db 1083 GTCCCCCGGACACGCGCTGGCGGGGTGGAGAGGCGACTTCCGACACATCCCTTCT 1142
Qy 1079 TCTCTAAATTTGACTGGAACAAATCGTAACTCTCTCCCGCTTCTGTTCCACCCCTCA 1138
Db 1143 TCTTGGCTTGAAGTGGATGCTCGGAGAGGCTGCGGCGGCTGCGGCGGCTTACCGGATTTG 1202
Qy 1139 AGTCTGACGATGACACTCCAAATTTGA 1166
Db 1203 AAGGTGCCACCGACACATGCAACTTGA 1230

RESULT 6

US-08-422-706B-12
; Sequence 12, Application US/08422706B
; Patent No. 5977333
; GENERAL INFORMATION:
; APPLICANT: Brook, J. David
; APPLICANT: Housman, David E.
; APPLICANT: Shaw, Duncan J.
; APPLICANT: Harley, Helen G.
; APPLICANT: Johnson, Keith J.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
; TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02713
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,706B
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,543
; FILING DATE: 08-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/023,612

; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/839,255
; FILING DATE: 20-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01545
; FILING DATE: 19-FEB-1993
; PRIOR APPLICATION DATA: PCT/GB93/00253
; FILING DATE: 05-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB9202485.0
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-5830A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-422-706B-12

Query Match 3.6%; Score 220.8; DB 2; Length 2726;
Best Local Similarity 53.7%; Pred. No. 2e-49;
Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;

Qy 194 CTCCTGTGATGAAGATTAGCAGCTGAGCACTTTGTCGGAAGATTCCGACACCATAG 253
Db 243 CCGAACTGGCCCCAGGACAAAGTAGCTGGCGGACTTTCTTGCACTGGGCGGACCATCGTGG 302
Qy 254 CTGAGTTACAGGAGCTCCAGCCCTTCGGCAAGACTTCGAACTTCAGAGTCTTTAGGTT 313
Db 303 TGAGGCTTAAGAGGTTCCGACTGCAGAGGAGCGACTTCGAGATTCTGAAGTGTACGAC 362
Qy 314 GTGGTCACTTTGCTGAAGTGCAGAGTGTGTAAGAGAGAAAGCAACCGGGAGCATCTATGCTA 373
Db 363 CGCGGGCGTTTACGAGAGGTAGCGGTAGTGAAGATGAAGCAGACGCGGCCAGGTGTATGCCA 422
Qy 374 TGAAGTGTATGAAGAAGAGGCTTTATTGSCCCAGGAGCAGGTTTCATTTTGGAGGAG 433
Db 423 TGAAGTATGAAGAAGTGGGACATGCTGAAGAGGGCGAGGTGTGCTCTCCGTGAGG 482
Qy 434 AGCGGAACATATTATCTCGAAGCAGACCCCGTGGATCCCAATTACAGTATGCTTTC 493
Db 483 AGAGGACCGTGTGTGTAATGGGACCGCGGTGGATCAGCAGCTGCACATTCGCTTCC 542
Qy 494 AGGACAAAATCACCCTTTATCTGATGAGGAAATATCAGCCTCGAGGGGACTTGTGTAC 553
Db 543 AGGATGAGAACTACCTGTACCTGGTCAATGAGTATTTACGTGGCGGGGACCTGCTGAC 602
Qy 554 TTTTGAATAGATATAGGACCACTTAGATGAAAACTGATACAGTTTACCTAGCTGAGC 613
Db 603 TGCTGAGCAAGTTTGGGAGCGGATTCGCGCGAGATGGCGCTTCTACCTGGCGGAGA 662
Qy 614 TGATTTTGGCTGTTCACAGCTTTCATCTGATGGGATAGTGCATCGAGACATCAAGCCTG 673
Db 663 TTGCTATGGCCATAGACTCGGTGACCGCTTGGCTAGTGTGCAAGGAGACATCAACCCG 722
Qy 674 AGAACATCTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTTTGGATCTGCGCGGA 733
Db 723 ACAACATCTCTGTTGACCGCACATCGGCTGGCGGACATCCGCTGGCGGACTTGGCTCTCA 782
Qy 734 AAATGAATTCACAAAGATGGTGAATGCCAAACTCCCGATTGGGACCCAGATTTACATGG 793
Db 783 AGCTGCGGCGAGATGGAACGGTGGCTGCTGGTGGCTGTGGGACCCAGACTACCTGT 842
Qy 794 CTCCTGAAGTGTGA---CTGTGATGAACGGGATGGAAGGACACCTACGCGCTTGGACT 850

843 CCCCAGATCTCTCAGGCTGTGGCGGTGGGCTGGGACAGGAGCTACGGGCCCGAGT 902
851 GTGACTGTGTGATGAGTGGGCTGATTCCTATGAGATGATTTATGGAGATCCCTCTCG 910
903 GTGACTGTGTGGCGCTGGGCTGATTCCTATGAGATGATTTATGGAGATCCCTCTCG 962
911 CAGAGGAACCTCTGCCAGAACCTTCAATACATTTATGAATTTCCAGCGGTTTGAAT 970
963 ACAGGATTCACAGCGGAGACCTATGGCAAGATCGTCCACTAAGAGGACCTCTCTC 1022
971 TTCAGATGACCCCAAGTACAGT---GACATTTCTGATCTGATCAAGTTTGT 1027
1023 TGCCCTGTGGAGGAGGCTCCCTGAGGAGGCTCGAGATTCATTCAGCGGTTCTGT 1082
1028 GCGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGC-----CATCTTTCT 1078
1083 GTCCCCCGAGACACGCTGGCGGGGTGGAGCAGCGGACTTCGGACACATCCCTTCT 1142
1079 TCTCTAAATTTGACTGGAACAACTTGTGTAATCTCTCTCCCTCTGTTCCACCTCA 1138
1143 TCTTGGCTCTGACTGGATGGTCTCGGAGACAGCGTGGCCCTTTACACCGGATTCG 1202
1139 AGTCTGACGATGACCTCCAAATTTGA 1166
1203 AAGGTGCCACGACACATGCACTTGA 1230

RESULT 7

US-08-422-699A-8
; Sequence 8, Application US/08422699A
; Patent No. 5955265
; GENERAL INFORMATION:
; APPLICANT: Brook, J. David
; APPLICANT: Housman, David E.
; APPLICANT: Shaw, Duncan J.
; APPLICANT: Harley, Helen G.
; APPLICANT: Johnson, Keith J.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
; TITLE OF INVENTION: DYSPROPHY GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02713
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,699A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/422,706
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/023,612
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/839,255
; FILING DATE: 20-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01545
; FILING DATE: 19-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00253
; FILING DATE: 05-FEB-1993
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-699A-8

Query Match 3.5%; Score 214; DB 2; Length 2511;
Best Local Similarity 54.2%; Pred. No. 1.4e-47;
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;

QY 244 GACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGCAAGAGCTTCGAAGTCAGAAGT 303
Db 43 GCCATCGTGTGAGGCTTAAGGAGGTCGAGTGCAGAGGAGCAGCTTCGAGATTTCTGAAG 102
QY 304 CTTGTAGTTGTGTCATCTTCTGAGTGCAGTGTGTAAGAGAGAAAGCAACCGGGGAC 363
Db 103 GTGATCGACCGCGGGGTTTCAGCGAGGTAGCGGTAGTGAAGATGAAGCAGACGCGGCCAG 162
QY 364 ATCTATGCTATGAAGTGTATGAAGAGAGAGGCTTTATTGGCCCGAGGAGAGGTTTCATTT 423
Db 163 GTGTATGCCATGAAGATCATGAACAAGTGGACATCTGAAGAGGGGCGAGGTGTCTGTC 222
QY 424 TTTGAGGAGAGCGGACATATTTATCTGAGACACAAGCCCGTGGATCCCCCAATTACAG 483
Db 223 TTCGTGAGGAGGAGGACGCTGTGGTGAATGGGACCGCGGTGGATCAGCAGCTGCAC 282
QY 484 TATGCTCTTTCAGGACAAATACCTTTATCTGATGGAGAAATATCAGCTTCGAGGGGAC 543
Db 283 TTCGCTTCCAGATGAGACTACCTGTACCTGGTCTATGGATTTTACGTGGCGGGGAC 342
QY 544 TTGCTGTACTTTTGAATAGATATGAGGACCAAGTTAGATGAAGAACTGATACAGTTTAC 603
Db 343 CTGCTGACACTGCTGAGCAAGTTTGGGAGCGGATTCGCGCGAGATGCGCGCTTCTAC 402
QY 604 CTAGCTGAGCTGATTTGGCTGTTCACAGCTTCTATGATGGATACCTGATCGAGAC 663
Db 403 CTGGCGGAGATTTCTATGCCCATAGACTCGGTGACCCGCTTGGCTTACGTGACAGGGAC 462
QY 664 ATCAAGCTTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTGGATTTGA 723
Db 463 ATCAAAACCGGACAAACATCTCTGTGGACCGCTGTGGCCACATCCGCTTGGCCGCTTCCG 522
QY 724 TCTGCGCGGAAATGAATTCAAACAAGATGTTGAATGCCAACTCCCGATTGGGACCCCA 783
Db 523 TCTTGTCTCAAGCTGGGCGAGATGAGAACGGTGGGTGCTGGTGGCTGTGGGCGACCCCA 582
QY 784 GATTACATGGCTCTGA---AGTGTGACTGTGATGAAGCGGGGATGGAAGGACCTTAC 840
Db 583 GACTACCTGTCCCGGAGATCTCGAGGCTGTGGCGGTGGGCGCTGGGACAGGACGCTAC 642
QY 841 GGCCTGGAATGACTGTGGTGTGAGTGGCGGTGATGCTATGAGATGATTTATGGGAGA 900
Db 643 GGGCCCGAGTGTGACTGGTGGCGGTGATTCGCTTATGAAATGTTCTATGGGAG 702
QY 901 TCCCCCTTCGAGAGGGGAACTTCGCCAGAACCTTCATTAATCAATTTGAAATTTCCAGCGG 960
Db 703 ACGCCTTCTACGCGGATTCACGCGGAGACCTTATGCAAGATCGTCCACTACAGGAG 762

QY 961 TTTTGAATTCACAGATGACCCCAAGTGAAGT---GACTTTCTTGTATCTGATTCAA 1017
Db 763 CACTCTCTCTGCGCGTGGGAGAGGGTCCCTGAGGAGCTCGAGACTTCATTCAG 822
QY 1018 AGCTTGTGTGGCGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTGC----- 1068
Db 823 CGGTGTGTGTCCTCCCGGAGACACGGCTGGCGGGGTGGAGCAGCGACTTCGGACA 882
QY 1069 CATCTTCTCTCTCTAAATAGTGAACAACATTGCTGAATCTCTCTCCCTTCGTT 1128
Db 883 CATCTTCTCTCTCTGCGCTGAGTGGTCTCGGAGACAGCTGCGCCCTTTTACA 942
QY 1129 CCCACCTCAAGTCTGACGATGACACCTCCAAATTTGA 1166
Db 943 CCGAATTCGAAGTGGCCCGGACACATGCACTTGA 980

RESULT 8

US-08-422-706B-8
; Sequence 8, Application US/08422706B
; Patent No. 5977333
; GENERAL INFORMATION:
; APPLICANT: Brook, J. David
; APPLICANT: Rousman, David E.
; APPLICANT: Shaw, Duncan J.
; APPLICANT: Harley, Helen G.
; APPLICANT: Johnson, Keith J.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,706B
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,543
; FILING DATE: 08-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/023,612
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/839,255
; FILING DATE: 20-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01545
; FILING DATE: 19-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00253
; FILING DATE: 05-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB9202485.0
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-5830A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:

LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-706B-8
Query Match
Best Local Similarity 54.2%; Pred. No. 1.4e-47;
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;
QY 244 GACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTGGGCAAGAGACTTCCAAAGTCAGAAGT 303
Db 43 GCCATCGTGTGAGGCTTAAGGAGGTCCTGAGGAGGACGACTTCGAGATTCTGAAG 102
QY 304 CTTGTAGGTTGTGTGTCACATTTTGTGAAGTGCAGGTTGAAGAGAGAAAGCAACCCGGGAC 363
Db 103 GTGATCGGACGCGGGCGTTACGAGAGTACGGTAGTGAAGATGAAGCAGACAGCGGCCAG 162
QY 364 ATCTATCTATGAAGTATGAAGAAAGAGGCTTTATTTGGCCCGAGAGCAGGTTTCATTT 423
Db 163 GTGTATGCCATGAAGATCATGAACAAAGTGGGACATGCTGAAGAGGGCGAGGTGTCGTG 222
QY 424 TTTGAGGAAGCGGCAACATATTATCTGAAGCACAAGCCCGTGGATCCCCCAATTACAG 483
Db 223 TTCCGTGAGGAGGAGCGTGTGTTGATGGGACCGCGGTGGATCACGACGCTGCAC 282
QY 484 TATGCCCTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGGGGAC 543
Db 283 TTCGCTTCCAGGATGAGAACTACCTGTACCTGGTTCATGGAGTATTACGTGGGCGGGAC 342
QY 544 TTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACTGTATACAGTTTAC 603
Db 343 CTGCTGACACTGCTGAGCAAGTTTGGGAGCGGATTCCGGCGGAGATGGCGCTTCTAC 402
QY 604 CTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTTCATCTGATGGGATACGTGATCAGAC 663
Db 403 CTGGCGGAGATTGTATGCCCATAGACTCGGTGCAACCGCTTGGCTTGGTGCACAGGGAC 462
QY 664 ATCAAGCCTGAGAACTTCTGTTGACCCACAGACACATCAAGCTGGTGGATTTGGA 723
Db 463 ATCAACCCGACACATCTGCTGACCGCTGTGGCCACATCCGCTGGCCGACTTCGGC 522
QY 724 TCTGCGCGGAAATGAATTCAAACAGATGGTGAATGCCAAACTCCGATTTGGGACCCCA 783
Db 523 TCTTGGCTCAAGCTGGCGGAGATGGAACCGGTGCGTGGTGGTGGTGGTGGTGGTGGTGG 582
QY 784 GATTACATGGCTCTCTGA---AGTGTGCTGATGATGAACGGGGATGGAAGGACCTTAC 840
Db 583 GACTACTGCTGCTCCCGAGATCTGAGGCTGTGGCGGTGGGCTGGGACAGGAGCTAC 642
QY 841 GGCTGAGCTGTGACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
Db 643 GGGCCCGAGTGTGACTGTGGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 702
QY 901 TCCCGCTTCGACAGGGAACCTCTGCGAAGCTTCAATAACATTATGAATTTCCAGCGG 960
Db 703 ACGCCCTTCTACCGGATTCACGCGGAGACTATGCAAGATCGTCCACTACAGAGAG 762
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGT---GACTTTCTTGTATCTGATTCAA 1017
Db 763 CACTCTCTCTGCGCTGGTGGACGAGGGTCCCTGAGGAGCTTCGAGACTTCATTCAG 822
QY 1018 AGCTTGTGTGGCGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTGC----- 1068
Db 823 CGGTGTGTGTCCTCCCGGAGACACGGTGGCGGGGTGGAGCAGCGACTTCGGACA 882
QY 1069 CATCTTCTCTCTCTAAATAGTGAACAACATTGCTGAATCTCTCTCCCTTCGTT 1128
Db 883 CATCTTCTCTCTCTGCGCTGAGTGGTCTCGGAGACAGCTGCGCCCTTTTACA 942

QY 1129 CCCACCTCAAGTGTGAGGATGACACCTCCAAATTTGA 1166
 DB 943 CCGGATTTCGAAGGTGCCACCGACACATGCAACTTCGA 980

RESULT 9

US-08-484-044-11
 ; Sequence 11, Application US/08484044
 ; GENERAL INFORMATION:
 ; APPLICANT: Caskey, C. T.
 ; APPLICANT: Fu, Ying-Hui
 ; APPLICANT: Friedman, David L.
 ; APPLICANT: Pizzuti, Antonio
 ; APPLICANT: Fenwick, Raymond G.
 ; TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESS: Fulbright & Jaworski, L.L.P.
 ; STREET: 1301 McKinney, Suite 5100
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: U.S.A.
 ; ZIP: 77010-3095
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patencin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,044
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/019,940
 ; FILING DATE: 19-FEB-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Paul, Thomas D.
 ; REGISTRATION NUMBER: 32,714
 ; REFERENCE/DOCKET NUMBER: D-5443
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 713/651-5325
 ; TELEFAX: 713/651-5246
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3182 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-484-044-11

Query Match 3.5%; Score 213.6; DB 1; Length 3182;
 Best Local Similarity 54.2%; Pred.No. 28-47;
 Matches 507; Conservative 0; Mismatches 414; Indels 15; Gaps 3;
 QY 246 CACCATAGCTAGTTACAGGAGCTCCAGCCTTCGGCAAAGGACTTCGAAGTCAGAAGTCT 305
 DB 743 CATCGTGGTGAAGGCTTAAGGAGGTCGACTGCAGAGGAGCAGCTTCGAGATTCGAAGGT 802
 QY 306 TGTAGTTGTGCTACTTGTCTGAAGTCAGGTGTTAGAGAGAAAGCAACCGGGGACAT 365
 DB 803 GATCGGACCGGGGGGTTTCAGCGAGGTAGCGGTAGTGAAGATGAAGCAGACGGGCCAGGT 862
 QY 366 CTATGCTATGAAGTGTCAAGAAAGGCTTTATTGGCCCCAGGAGCAGGTTTCATTTTT 425
 DB 863 GTATGCCATGAAGATCATGAACAAGTGGACATGCTGNAAGGGGCGGAGGTGTCGCTT 922
 QY 426 TGAGAGAGCGGCAACATATTATCTGAGCACAAGCCCGTGGATCCCCCAATTTACAGTA 485
 DB 923 CCGTGAGGAGGAGCGGTGTTGGTGAATGGGGACCGCGGTGGATCATCAGCAGTGCACCT 982

QY 486 TGCCTTTCAGGACAAAAATCACCTTTATCTGTATGGAGGAATATCAGCCTCGAGGGGACTT 545
 DB 983 CGCCTTCAGGATGAGAACTACCTGTACTGTGTCATGGAGTATTACGTGGGGGGGACCT 1042
 QY 546 GCTGTACCTTTTGAATAGATATGAGGACCAAGTATGAAATGAAACCTGATACAGTTTACCT 605
 DB 1043 GCTGACACTGTCTGAGCAAGTTTGGGGAGCGGATTCCGCCGAGATGGCGCTCTTACCT 1102
 QY 606 AGCTGAGCTGATTTTGGCTGTTCAAGCGTTTCATCTGATGGGATACGTGCTACAGACAT 665
 DB 1103 GCGGAGATTGTCTATGGCCATAGACTCGGTGCACCGGCTTGGCTACGTGTCACAGGGACAT 1162
 QY 666 CAAGCTCAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTTGGATC 725
 DB 1163 CAACCCGACAACATCTCTGTGGACCGCTGTGGCCACATCCGCTGGCCGACTTGGGCTC 1222
 QY 726 TGCCGCGAAAAATGAATCAACAAGATGGTGAATGCCAAACTCCCGATTGGGACCCGAGA 785
 DB 1223 TTGCTCTAAGCTGCGGGCAGATGGAACGTCGCTGCTGTGGCTGTGGGACCCGAGA 1282
 QY 786 TTACATGCTCCTGA---AGTGTGACTGTGATCAACGGGATCGAAAGGACCTACGG 842
 DB 1283 CTACTGTCCCCGAGATCCTGCAAGCTGTGGGCGGTGGGCTGGGACAGGACGCTACGG 1342
 QY 843 CTGACTGTGACTGGTGGTCAAGTGGCGGTGATTCCTATGAGATGATTTATGGGAGATC 902
 DB 1343 GCGGAGTGTACTGTGGCGCTGGGTGATTCGCTATGAAATGTTCTATGGGACAGAC 1402
 QY 903 CCCCTTCGAGAGGAACTCTGCGACAACCTCAATAACATTTATGATTTCCAGCGGTT 962
 DB 1403 GCGCTTCTACGCGATTCACGCGGAGACATATGGCAAGATCGTCCACTACAGGAGCA 1462
 QY 963 TTTGAAATTTCCAGATGACCCCAAGTGACAGT---GACTTTCTTGTATGATTTCAAAG 1019
 DB 1463 CCTCTCTGCTGCTGGTGGACGAAGGGTCCCTGAGGAGGCTCGAGACTTCATTACGG 1522
 QY 1020 CTTGTTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGC-----CA 1070
 DB 1523 GTTGTCTGTCCCGGAGACAGCGCTGGCGCGGTGGAGCAGGAGCTTCGGSACACA 1582
 QY 1071 TCCTTTCTTCTTAAATTTGACTGGAAACAACATTCGTAATCTCTCTCCCGGACAGGCTGCTTCC 1130
 DB 1583 TCCTTTCTTCTTGGCTGCTGACTGGGATGCTCTCCGGACAGGCTGCTGCTTACAC 1642
 QY 1131 CACCTCAAGTGTGAGATGACACCTCCCAATTTGA 1166
 DB 1643 GGATTTGGAAGGTGCCACCGACACATGCAACTTCGA 1678

RESULT 10

US-09-804-471A-3
 ; Sequence 3, Application US/09804471A
 ; Patent No. 6479269
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marion et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001164
 ; CURRENT APPLICATION NUMBER: US/09/804,471A
 ; CURRENT FILING DATE: 2001-03-13
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 174493
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(174493)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-804-471A-3

Query Match 3.3%; Score 205; DB 4; Length 174493;
Best Local Similarity 100.0%; Pred. No. 5.7e-44;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 GGTGAATGCCAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 812
DB 130289 GGTGAATGCCAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 130348

QY 813 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGGTCTCAGTGGGCGT 872
DB 130349 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGGTCTCAGTGGGCGT 130408

QY 873 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 932
DB 130409 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 130468

QY 933 CTTCAATACATTATGAATTTCCAG 957
DB 130469 CTTCAATACATTATGAATTTCCAG 130493

RESULT 11
US-10-238-709-3
; Sequence 3, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROF
; FILE REFERENCE: CLO01164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(174493)
; OTHER INFORMATION: n = A,T,C or G

US-10-238-709-3

Query Match 3.3%; Score 205; DB 4; Length 174493;
Best Local Similarity 100.0%; Pred. No. 5.7e-44;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 GGTGAATGCCAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 812
DB 130289 GGTGAATGCCAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 130348

QY 813 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGGTCTCAGTGGGCGT 872
DB 130349 GATGAACGGGGATGAAAGGACCTACGGCTGGACTGTGACTGGTCTCAGTGGGCGT 130408

QY 873 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 932
DB 130409 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 130468

QY 933 CTTCAATACATTATGAATTTCCAG 957
DB 130469 CTTCAATACATTATGAATTTCCAG 130493

RESULT 12
US-08-630-822A-61
; Sequence 61, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.

APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDIA
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,822A
FILING DATE: 11-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 2706 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 5..2706
US-08-630-822A-61

Query Match 3.1%; Score 189.2; DB 2; Length 2706;
Best Local Similarity 54.1%; Pred. No. 7.9e-41;
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;

QY 281 CAAAGGACTTCGAAAGTCAGAGTCTTGTAGTGTGCTGCTCACTTTTGTCTGAAAGTGCAGGTGG 340
DB 78 CAGATGATTTAATTAAATAAAGTATTGTCGAGGAGCATTTGGTGAAGTACAGTTAG 137

QY 341 TAAGAGAGAAAGCAACCGGGACATCTATGCTATGAAAGTGAAGAGAGAGCTTTAT 400
DB 138 TCGACACAAATCAACTGCACAGTTTTTGTATGAAACGCCTATCAAAATTTGAAATGA 197

QY 401 TGGCCAGGACGAGTTTCATTTTTTGAGGAAGAGCGGAACATATTATCTCGAAGCACA 460
DB 198 TTAAGAGACAGACTCTGCATTTTTTTGGGAGAGACGTCATATAATGCTCATCAAAAT 257

QY 461 GCCCGTGGATCCCCAATTACAGTATGCTTTTTCAGGACAAAAATCACTTTATCTGATGG 520
DB 258 CAGAATGGATTGTACAATTACATTTTGTCTTTTCAAGATCAAAAATATCTTTATATGTC 317

QY 521 AGGAATATCAGCTGGAGGGACCTGCTGACATTTTGAATAGATATGAGGACCAAGTTAG 580
DB 318 TGGATTATATCCGGGGGGTGCATTTGTTGAGTCTTATG-----TCGATTATGAATTC 371

QY 581 ATGAAACCTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTTTCACAGCGTTCA 640
DB 372 CAGAAAATGGCAATGTTCTATAAATGGAAGTGGTGTAGCACTTGATCAATTCAT 431

QY 641 TGATGGGATAGCTGATCGAGACATCAAGCTCGAGACATTTCTGTTGACCGCAGGAC 700
DB 432 CCATGGGATTTGTACATCGTGATGTTAAACCTGATATATGCTTTCTAGACAAATATGCTC 491

QY 701 ACATCAAGCTGGTGAATTTTGGATCTGCGCGGAAAAATGAATTCAACAAAGATGTTGAATG 760

Db	492	ATTTAAAGTTAGCTGACTTTGGAACTTGATGAANAATGGATACAGATGGTTTGGTAGCTT	551
Qy	761	CCAAACTCCCGATTGGGACCCGAGATTACATGGCTCTGAAGTGCTGACTGTGATGAACG	820
Db	552	CTAATAATGTGTGTGGAAACCGCTGATTACATTTCTCCGAAGT-----TTTGCAGTCCC	605
Qy	821	GGGATGAAAAGGCAACCTACGGCTCGACATGTGACTGGTGGTCACTGAGTGGCGGTGATTCGCT	880
Db	606	AAGTGTGTGAAGGAGTTTACGGTCTGTGAATGCCGATTTGGTGGTCTGTGGGAATTTTGT	665
Qy	881	ATGAGATGATTTATGGAGATCCCCCTTCGAGAGGGAACTCTGCGCAGAACCTTCAATA	940
Db	666	ATGAAATGTTATTGGAGAAACACCTTTTATGCAGACAGTTTGGTTGGAACTTACAGTA	725
Qy	941	ACATTATGAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTCAGCAGTGACT	1000
Db	726	AAATTTATGGATCACAGAAACTCATTAACTTTTCTCCAGAAAGTGGAAATTAAGCCAATATG	785
Qy	1001	TTCTTGATCTGATTCAAAGCTTGTGTGCGCCAGAAAGAG-----AGACTGAAGT	1051
Db	786	CCCGATCTTTGATACAAAGGATTTTAAACAGACAGAACACAGCGTTTAGGCAGAAATCAAG	845
Qy	1052	TTGAAGTCTTTGCTGCGCATCTTTCTTCCTCTAAAATTTGACTGGA-----ACAACATTC	1105
Db	846	TGGAAGAAATTAACAGACATCCATTTTTCATAATGATCAATGACTTTTGTGACAAATTTAA	905
Qy	1106	GTAACTCTCTCCGCCCTTGTGCCACCTCAAGTCTGACGATGACACCTCAATTTTG	1165
Db	906	GAGACTCTGCCCACTGTAGTCCAGAGCTGATGGTGTGATGATGATACAAGGAATTTG	965
Qy	1166	ATGA 1169	
Db	966	ATGA 969	

RESULT 13

US-09-005-069-61
Sequence 61, Application US/09005069
Patent No. 5932470
GENERAL INFORMATION:
APPLICANT: FRANK, GLENN R.
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDA
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,069
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,622
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223

RESULT 15
 US-09-004-730A-20
 ; Sequence 20, Application US/09004730A
 ; Patent No. 6485968
 ; GENETAL INFORMATION:
 ; APPLICANT: Weber, Eric
 ; APPLICANT: Wu Hunter, Shirley
 ; APPLICANT: Sim, Gek-kee
 ; APPLICANT: Frank, Glenn
 ; APPLICANT: Wallenfels, Lynda
 ; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO C
 ; FILE REFERENCE: 2618-17-C5-PUS-1
 ; CURRENT APPLICATION NUMBER: US/09/004,730A
 ; CURRENT FILING DATE: 1998-01-08
 ; PRIOR APPLICATION NUMBER: PCT/97US/18669
 ; PRIOR FILING DATE: 1997-10-15
 ; NUMBER OF SEQ ID NOS: 150
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 2706
 ; TYPE: DNA
 ; ORGANISM: Ctenocephalides felis

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 10:49:43 ; Search time 1455 Seconds

(without alignments)
17982.579 Million cell updates/sec

Title: US-10-017-216-3

Perfect score: 6159
Sequence: 1 atgttgagttcaaatagg.....ttctgagaacagattattgc 6159

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6159	100.0	6574	6	AAD39191 Human MDP
2	5668	92.0	8603	8	AAL55215 Human CRI
3	5666.6	92.0	6156	8	AAL55217 Human CRI
4	5666.4	92.0	6298	6	AAD38864 Human kin
5	5666.2	92.0	6165	8	AAL55214 Human CRI
6	5661.4	91.9	6165	6	ABQ78870 Human kin
7	5650.4	91.7	6159	4	AAS06701 Polynucle
8	5631	91.4	6189	6	ABS63436 RHO/RAC-1
9	5631	91.4	6189	6	ADA05641 Human NOV
10	5629.8	91.4	6201	6	ABS63435 Human cdn
11	5629.8	91.4	6201	7	ADA05653 Human NOV
12	5358	87.0	5877	6	ABQ78871 Human kin
13	3475.2	56.4	6609	3	ACQ77568 Human ORF
14	2475.2	41.8	2896	5	ABV30132 Human pro
15	2564.8	41.6	3131	4	ABA08361 Human RHO
16	2436.8	39.6	5261	8	AAL55216 Human CRI
17	2426.8	39.4	5251	9	ADD89966 Human can
18	2414	39.2	2542	7	ADA05647 Human NOV
19	2317.2	37.6	2437	7	ADA05645 Human NOV
20	2126.8	34.5	2693	9	ADE09823 Novel DNA
21	1841	29.9	1870	7	ADA05643 Human NOV
22	1524.2	24.7	1870	7	ADA05649 Human NOV
23	1467.6	23.8	1915	7	ADA05651 Human NOV

24	1395	22.6	1485	7	ABZ68725	Abz68725 Nucleotid
25	1395	22.6	1745	7	ABZ68726	Abz68726 Nucleotid
26	1393.4	22.6	2086	6	AAD26454	Aad26454 Human kin
27	1391.8	22.6	1515	7	ACA61394	Aca61394 cDNA enco
28	1391.8	22.6	1515	9	AAD59938	Aad59938 Human kin
29	1082.4	17.6	2380	7	ABZ68776	Abz68776 Nucleotid
30	952.4	15.5	2162	7	ABX71191	Abx71191 Novel hum
31	930	15.1	1048	5	AAS79753	Aas79753 DNA encod
32	701.4	11.4	995	4	ABA08479	Aba08479 Human cit
33	648	10.5	1058	4	AAP22502	Aaf22502 Human bre
34	592	9.6	817	4	AAP22603	Aaf22603 Human bre
35	443.4	7.2	446	5	ABV15823	Abv15823 Human pro
36	436	7.1	485	5	ABV45624	Abv45624 Human pro
37	433.2	7.0	580	4	AAI92420	Aai92420 Human pol
38	350.4	5.7	354	2	AAV87531	Aav87531 EST clone
39	337	5.5	396	4	AAH99003	Aah99003 Marine ES
40	258	4.2	258	7	ACA55915	Aca55915 Mouse sig
41	250	4.1	254	2	AAV89320	Aav89320 EST clone
42	244	4.0	3835	7	ABT33346	Abt33346 NOX DNA
43	244	4.0	3985	7	ABT33347	Abt33347 NOX DNA
44	243.6	4.0	1530	6	AAD38845	Aad38845 Human kin
45	243.6	4.0	5373	6	AAD30567	Aad30567 Human kin

ALIGNMENTS

RESULT 1
AAD39191
ID AAD39191 standard; cDNA; 6574 BP.

XX AC AAD39191;

XX DT 04-OCT-2002 (first entry)

XX DE Human MDPK cDNA.

XX KW Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;
tumorigenesis; tumour growth; tumour metastasis; viral infection;
skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;
immune disorder; neoplastic disorder; gene therapy; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT 5'UTR 1..18

XX FT CDS /tag= a

XX FT CDS 19..6180

XX FT CDS /tag= b

XX FT misc_feature /product= "Human MDPK protein"

XX FT misc_feature 19..6177

XX FT /tag= c

XX FT /note= "This region is specifically referred in claim 1

as SEQ ID NO:3"

XX 3'UTR 6181..6574

XX /tag= d

XX WO200234896-A2.

XX PD 02-MAY-2002.

XX PF 23-OCT-2001; 2001WO-US050636.

XX PR 23-OCT-2000; 2000US-0242429P.

XX XX (MILL-) MILLENNIUM PHARM INC.

XX PI Kapeller-Libermann R;

XX XX WPI; 2002-479720/51.

XX DR P-FSDB; AAE24079.

XX XX Human myotonic dystrophy type protein kinase polypeptide and

Db 1819 CGGAAGCGACAGAATGTGAGATAAACTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAA 1878
QY 1861 GTGGAGAAATATGCGAAATCGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
Db 1879 GTGGAGAAATATGCGAAATCGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938
QY 1921 CTCGAAGAAACTGGAAGAGCTGCAAGAGAGCGAGAGCGGAGAGGAGCTGGAAGCTG 1980
Db 1939 CTCGAAGAAACTGGAAGAGCTGCAAGAGAGCGAGAGCGGAGAGGAGCTGGAAGAGCTG 1998
QY 1981 CAGAACCGAGAGAGATTTCTTCTCAAGGCATCAAGAAAGAGCTGGTGAAGAGCTGAGGAAGC 2040
Db 1999 CAGAACCGAGAGAGATTTCTTCTGAAGCATCAGAAAGAGCTGGTGAAGAGCTGAGGAAGC 2058
QY 2041 CCGCATTTCTCGGAACAAAGTAAAGAGACTAGAGACCATGGAAGCGTAGAGAAACAGA 2100
Db 2059 CGCCATTTCTCGGAACAAAGTAAAGAGACTAGAGACCATGGAAGCGTAGAGAAACAGA 2118
QY 2101 CTGAAGGATGATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2160
Db 2119 CTGAAGGATGATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2178
QY 2161 GAGCTCGAAGAGAAACATCGGAGGCCCCAAGTCTCAGCCCCAGACCTAGAAGTGACCTG 2220
Db 2179 GAGCTCGAAGAGAAACATCGGAGGCCCCAAGTCTCAGCCCCAGACCTAGAAGTGACCTG 2238
QY 2221 AAACAGAAAGAGAGAGCTATAGGAAGAAAGATTAAGTGTGGAACAATCAGATAAAGAAA 2280
Db 2239 AAACAGAAAGAGAGAGCTATAGGAAGAAAGATTAAGTGTGGAACAATCAGATAAAGAAA 2298
QY 2281 GACCTGGCTGCAAGAGAGACATGGAGAACATGATGCAGAGACACAGAGAGAGGCCCAT 2340
Db 2299 GACCTGGCTGCAAGAGAGACATGGAGAACATGATGCAGAGACACAGAGAGAGGCCCAT 2358
QY 2341 GAGAAAGGCAAAATTTCTACGGAACACAGAGGCGATGATCAATGCTATGATTTCCAGATC 2400
Db 2359 GAGAAAGGCAAAATTTCTACGGAACACAGAGGCGATGATCAATGCTATGATTTCCAGATC 2418
QY 2401 AGATCCCTGGACAGAGATTTGTGNACTGTCTGAGGCCAATAAATTTGCGACCAATAGC 2460
Db 2419 AGATCCCTGGACAGAGATTTGTGNACTGTCTGAGGCCAATAAATTTGCGACCAATAGC 2478
QY 2461 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGAGATGATTTCTGAACTCAGGCAA 2520
Db 2479 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2538
QY 2521 CAGAAATTTTACTGTGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAATACTGGAG 2580
Db 2539 CAGAAATTTTACTGTGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAATACTGGAG 2598
QY 2581 GAGCAGCTGGAGAAAGATCAGCCACCAAGACACAGTGACAAGATTCGGGTGCTGGAACCTG 2640
Db 2599 GAGCAGCTGGAGAAAGATCAGCCACCAAGACACAGTGACAAGATTCGGGTGCTGGAACCTG 2658
QY 2641 GAGACAAGATTTGGGAGGTCAGTCTAGAGCACGAGGACAGAAACTGGAGCTCAAGGCG 2700
Db 2659 GAGACAAGATTTGGGAGGTCAGTCTAGAGCACGAGGACAGAAACTGGAGCTCAAGGCG 2718
QY 2701 CAGCTCACAGAGCTACAGCTCTCCCTGCAAGGCGGAGTCAAGTTGACGCCCTGCGAG 2760
Db 2719 CAGCTCACAGAGCTACAGCTCTCCCTGCAAGGCGGAGTCAAGTTGACGCCCTGCGAG 2778
QY 2761 GCTGCAAGGCGGCCCTGGAGAGCCAGCTTCGCGAGCGGAAGACAGAGCTGGAAGAGACC 2820
Db 2779 GCTGCAAGGCGGCCCTGGAGAGCCAGCTTCGCGAGCGGAAGACAGAGCTGGAAGAGACC 2838
QY 2821 ACAGCAAGAGCTGAAAGAGAGATCCAGGCACTCAGCGGCACTAGAGATGAATTCAGGCG 2880
Db 2839 ACAGCAAGAGCTGAAAGAGAGATCCAGGCACTCAGCGGCACTAGAGATGAATTCAGGCG 2898
QY 2881 AAATTTGATGCTCTTGTAAACAGCTGTACTGTAAATCACAGACTGGAGGAGCAGCTAAAC 2940
Db 2899 AAATTTGATGCTCTTGTAAACAGCTGTACTGTAAATCACAGACTGGAGGAGCAGCTAAAC 2958

QY 2941 CAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAAACCTTCTACTTGTCCAAACAATC 3000
Db 2959 CAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAAACCTTCTACTTGTCCAAACAATC 3018
QY 3001 GATGAGGCTTCTGGCGCAACGACGAGATGTACAACCTGCGAAGTGAAAGTGGACCATCTC 3060
Db 3019 GATGAGGCTTCTGGCGCAACGACGAGATGTACAACCTGCGAAGTGAAAGTGGACCATCTC 3078
QY 3061 CGCCGGGAGATCAGGAACCGAGAGATGACCTTTACAGCCAGAGCAAAACGATGGAGGCT 3120
Db 3079 CGCCGGGAGATCAGGAACCGAGAGATGACCTTTACAGCCAGAGCAAAACGATGGAGGCT 3138
QY 3121 CTGAAGACAGCTGCACCATCTCTGGAGAACAGGTCAATGATTTGGAGGCCCTTAAACGAT 3180
Db 3139 CTGAAGACAGCTGCACCATCTCTGGAGAACAGGTCAATGATTTGGAGGCCCTTAAACGAT 3198
QY 3181 GAGCTGCTAGAAAAAGAGCGCAGTGGAGGCCCTGGAGGAGCGTCTCTGGGTGATGAGAAA 3240
Db 3199 GAGCTGCTAGAAAAAGAGCGCAGTGGAGGCCCTGGAGGAGCGTCTCTGGGTGATGAGAAA 3258
QY 3241 TCCAGATTTGAGTGTCTGGGTTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC 3300
Db 3259 TCCAGATTTGAGTGTCTGGGTTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC 3318
QY 3301 AGGCGGAGAGCGGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3360
Db 3319 AGGCGGAGAGCGGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3378
QY 3361 GAGCACAGGCTGAGATTTCTGGTCTCCAGAGGCTCTCAAAGAGCAGAGCTGGAAGGCC 3420
Db 3379 GAGCACAGGCTGAGATTTCTGGTCTCCAGAGGCTCTCAAAGAGCAGAGCTGGAAGGCC 3438
QY 3421 GAGAGCTCTCTGCAAGACTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAAT 3480
Db 3439 GAGAGCTCTCTGCAAGACTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAAT 3498
QY 3481 GCCCGAAGCTTTACAGCAGAAAGCTGGAGACTCAACGAGAGCTCAAAACAGAGGCTTCTCGAA 3540
Db 3499 GCCCGAAGCTTTACAGCAGAAAGCTGGAGACTCAACGAGAGCTCAAAACAGAGGCTTCTCGAA 3558
QY 3541 GAGCAAGCCTTTACAGCAGCAGATGGAGCTGCAGAAAAATCAATTTCCGCTCTGACT 3600
Db 3559 GAGCAAGCCTTTACAGCAGCAGATGGAGCTGCAGAAAAATCAATTTCCGCTCTGACT 3618
QY 3601 CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660
Db 3619 CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3678
QY 3661 GAGTATCAGCTGGAAGAAACATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGC 3720
Db 3679 GAGTATCAGCTGGAAGAAACATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGC 3738
QY 3721 ACTATTTCTCAACAAACCAACTCATTGATTTTCTGCAAGCCAAATGGACCAACCTGCT 3780
Db 3739 ACTATTTCTCAACAAACCAACTCATTGATTTTCTGCAAGCCAAATGGACCAACCTGCT 3798
QY 3781 AAAAAGAAAAAGGCTTTATTTAGTTCGACGGAAGAGGACCTGCTTTACCCACACAGGTT 3840
Db 3799 AAAAAGAAAAAGGCTTTATTTAGTTCGACGGAAGAGGACCTGCTTTACCCACACAGGTT 3858
QY 3841 CCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGCAAG 3900
Db 3859 CCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGCAAG 3918
QY 3901 CTAGAGGAAGGCCCTTCAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCC 3960
Db 3919 CTAGAGGAAGGCCCTTCAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCC 3978
QY 3961 CACCGCAAGCAACGGAACCCACACCCATCCAGCTCAGCCACCGCAGGAGCAGCAGATC 4020
Db 3979 CACCGCAAGCAACGGAACCCACACCCATCCAGCTCAGCCACCGCAGGAGCAGCAGATC 4038

AA155215	AA155215 standard; DNA; 8603 BP.	
XX	AA155215;	
XX	01-MAY-2003 (first entry)	
XX	Human CR1K related DNA sequence, SEQ ID No 4.	
XX	Anorectic; hypotensive; cardiant; antilipaeamic; cerebroprotective;	
XX	antigout; osteopathic; antiarthritic; cytostatic; antidepressant;	
KW	immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;	
KW	neuroprotective; antiinflammatory; antidiabetic; analgesic;	
KW	human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;	
KW	obesity; comorbidity; cancer; anorexia; cachexia; bulimia;	
KW	central nervous system disorder; chronic obstructive pulmonary disease;	
KW	diabetes; pain; ds.	
XX	Homo sapiens.	
OS	WO2003/004523-A1.	
PN	16-JAN-2003.	
XX	28-JUN-2002; 2002WO-BP007156.	
XX	02-JUL-2001; 2001US-0301841P.	
PR	11-DEC-2001; 2001US-0338651P.	
PR	25-APR-2002; 2002US-0375014P.	
XX	(PARB) BAYER AG.	
XX	Zhu Z;	
PI	WPI; 2003-221576/21.	
XX	New human citron rho/rac-interacting kinase (CR1K) polypeptide and	
PT	polynucleotide, useful in preventing, ameliorating or treating diseases	
PT	associated with human CR1K dysfunction, e.g. obesity, diabetes or	
PT	Alzheimer's disease.	
XX	Disclosure; Fig 4; 237pp; English.	
PS	The invention relates to an isolated polynucleotide encoding a human	
CC	citron rho/rac-interacting kinase polypeptide. The isolated	
CC	polynucleotide comprises a 6165 or 8603 base pair sequence, given in the	
CC	specification. The human citron rho/rac-interacting kinase (CR1K)	
CC	polypeptide and polynucleotide are useful in preventing, ameliorating, or	
CC	treating diseases associated with human CR1K dysfunction such as obesity	
CC	and obesity-associated comorbidities (e.g. hypertension, coronary artery	
CC	disease, hyperlipidemia, stroke, gout, osteoarthritis, some types of	
CC	cancer including endometrial, breast, prostate and colon cancer),	
CC	anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood	
CC	disorders, anxiety disorders, Parkinson's disease or Alzheimer's	
CC	disease), chronic obstructive pulmonary disease, or diabetes. These can	
CC	also be used to treat pain associated with the disorders. The human CR1K	
CC	polypeptide is also useful in diagnostic assays or in genetic testing.	
CC	The expression vector or the reagent is useful in preparing a medicament	
CC	for modulating the activity of a human CR1K in a disease, e.g. obesity, a	
CC	central nervous system disorder, or chronic obstructive pulmonary	
CC	disease. The fusion protein is useful for generating antibodies against a	
CC	CR1K polypeptide and for use in various assay systems. The methods are	
CC	useful in producing and detecting the polynucleotide and polypeptide and	
CC	in screening for agents that modulate the activity of the human CR1K	
CC	polypeptide. This polynucleotide sequence represents a DNA sequence	
CC	relating to the human CR1K protein of the invention	
XX	Sequence 8603 BP; 2305 A; 2206 C; 2215 G; 1877 T; 0 U; 0 Other;	
XX	Query Match 92.0%; Score 568; DB 8; Length 8603;	
XX	Best Local Similarity 95.5%; Pred. No. 0;	
XX	Matches 5962; Conservative 0; Mismatches 35; Indels 243; Gaps 3;	

1081 TCTAAAATTGACTGGAACAACTTCGTAATCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1140
1141 TCTGACGATCAGACCTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT 1200
1141 TCTGACGATGACACCTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT 1200
1201 CCGTGCAGCTGAGCCCTCAGCCTTCTCGGTGAAAGAACTGCGCTTTGTGGGTTTCG 1260
1201 CCGTGCAGCTGAGCCCTCAGCCTTCTCGGTGAAAGAACTGCGCTTTGTGGGTTTCG 1260
1261 TAGCAGAAGCACTGGGATTCCTGTAGATCTGAGTCTGTTGTGTCGGGTCTGACTCC 1320
1261 TAGCAGAAGCACTGGGATTCCTGTAGATCTGAGTCTGTTGTGTCGGGTCTGACTCC 1320
1321 CCGTGCAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
1321 CCGTGCAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
1381 TCTCAGGACAACTGTCACAGATGGAGCAGGAATGACCCGGTTACATCGGAGTGTCA 1440
1381 TCTCAGGACAACTGTCACAGATGGAGCAGGAATGACCCGGTTACATCGGAGTGTCA 1440
1441 GAGGTGGAGCTGTGCTTGTAGTCAGAGGAGGTGGAGCTCAAGCCCTCTCAGACTCAGAGA 1500
1441 GAGGTGGAGCTGTGCTTGTAGTCAGAGGAGGTGGAGCTCAAGCCCTCTCAGACTCAGAGA 1500
1501 TCCCTCTCTGAGCAGGACCTTGCTACTCATCACAAGATGCAATGAGTCTTAAAGCAAGT 1560
1501 TCCCTCTCTGAGCAGGACCTTGCTACTCATCACAAGATGCAATGAGTCTTAAAGCAAGT 1560
1561 TTGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
1561 TTGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
1621 GATATCAGAGAGCAGGCGGAAGCTCCAAAGAAATCAAAAGAGCAGGAGTACCAGGCTCAA 1680
1621 GATATCAGAGAGCAGGCGGAAGCTCCAAAGAAATCAAAAGAGCAGGAGTACCAGGCTCAA 1680
1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAGAGGATCTGCTCAGCAGAGA 1740
1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAGAGGATCTGCTCAGCAGAGA 1740
1741 CGGAGTGTCTCTACGAACTCGAGCTGAGAGAGTCTCGGCTTGTCTGCTGCTGCTGCTCAAG 1800
1741 CGGAGTGTCTCTACGAACTCGAGCTGAGAGAGTCTCGGCTTGTCTGCTGCTGCTGCTCAAG 1800
1801 CGGAAAGCGAGAGATGTGAGCTAACTGTTGAGGCTTAAGGATCAAGGAGGCTGAA 1860
1801 CGGAAAGCGAGAGATGTGAGCTAACTGTTGAGGCTTAAGGATCAAGGAGGCTGAA 1860
1861 GTGGAAGATATCGGAACTGGAGAGAGTCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
1861 GTGGAAGATATCGGAACTGGAGAGAGTCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920
1921 CTCACAGAGAACTGGAGAGGCTG----- 1945
1921 CTCACAGAGAACTGGAGAGGCTG----- 1945
1946 -----CAAGAGCGAGCGGAGGAGCTGGAGAGCTGGAGAGCTGAGAGCGGAG 1992
1946 -----CAAGAGCGAGCGGAGGAGCTGGAGAGCTGGAGAGCTGAGAGCGGAG 1992
1981 AATATCGGCGAGCAAGAGCGAGCGGAGGAGCTGGAGAGCTGGAGAGCTGAGAGCGGAG 2040
1993 GATTTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAAGCGGCGCATTTCTG 2052
2041 GATTTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAAGCGGCGCATTTCTG 2100
2053 GAGAAACAGGTAGAGAGCTAGAGAGCTGGAGGCTGAGAGAGAGAGAGTGAAGTATGAC 2112
2101 GAGAAACAGGTAGAGAGCTAGAGAGCTGGAGGCTGAGAGAGAGAGAGTGAAGTATGAC 2160
2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2172

2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2220
2173 AAAACATCGGAGGCGCCCAAGTCTCAGCCAGCACCTAGAAAGTGCACTCTGAAACAGAAAGAG 2232
2221 AAAACATCGGAGGCGCCCAAGTCTCAGCCAGCACCTAGAAAGTGCACTCTGAAACAGAAAGAG 2280
2233 CAGCACTATGAGGAGAAAGATTTAAAGTGTGGCAATCAGATTAAGAAAGAGCTGGCTGAC 2292
2281 CAGCACTATGAGGAGAAAGATTTAAAGTGTGGCAATCAGATTAAGAAAGAGCTGGCTGAC 2340
2293 AAGGAGACACTGGAGAACATGATGATCAGAGACACAGAGAGAGGCGCCATGAGAAAGGCAAA 2352
2341 AAGGAGACACTGGAGAACATGATGATCAGAGACACAGAGAGAGGCGCCATGAGAAAGGCAAA 2400
2353 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2412
2401 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2460
2413 CAGAGGATTTGGAACTGTCTGAGCCCAATAAATTCAGCAAAATAGCAGTCTTTTACC 2472
2461 CAGAGGATTTGGAACTGTCTGAGCCCAATAAATTCAGCAAAATAGCAGTCTTTTACC 2520
2473 CAAAGGAACTGAAAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTAC 2532
2521 CAAAGGAACTGAAAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTAC 2580
2533 CTGAGAGACAGGCTGGAGAGTTGGAGCCCAAGAACCGAACTGGAGAGAGCTGGAG 2592
2581 CTGAGAGACAGGCTGGAGAGTTGGAGCCCAAGAACCGAACTGGAGAGAGCTGGAG 2640
2593 AAGATCAGCCACCAAGACCAAGTGAAGAACTGGCTGCTGGAACTGGAGAGCAAGATTG 2652
2641 AAGATCAGCCACCAAGACCAAGTGAAGAACTGGCTGCTGGAACTGGAGAGCAAGATTG 2700
2653 CGGAGGTCAGTCTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCCAGCTCAGAG 2712
2701 CGGAGGTCAGTCTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCCAGCTCAGAG 2760
2713 CTACAGTCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCTTCAAGGCTGACCGGCG 2772
2761 CTACAGTCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCTTCAAGGCTGACCGGCG 2820
2773 GCTTGGAGAGCAGCTTCCGAGGCGGAGCAGAGCTGAGAGCTGAGAGAGCAGAGAGCT 2832
2821 GCTTGGAGAGCAGCTTCCGAGGCGGAGCAGAGCTGAGAGAGCAGAGAGCAGAGAGCT 2880
2833 GAGAGAGATCCAGGAGCTCAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 2892
2881 GAGAGAGATCCAGGAGCTCAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 2940
2893 CTTCTGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGCT 2952
2941 CTTCTGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGCT 3000
2953 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3012
3001 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3060
3013 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3072
3061 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3120
3073 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3132
3121 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3180
3133 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3192
3181 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3240
3193 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3252
3241 GAGAGAGCTGCTACTGTAATCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3300

Db	5416	GCCTCTTCCAAACAGCTTCCCTTGCTCAATCGTGCAGGTGAACACGCGAGGGCAGCGAGAG	5475
Qy	5323	GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGACGTAGC	5382
Db	5476	GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAGACGTAGC	5535
Qy	5383	CGCA CAGACGATCTCAAGTGGAGTCGTTTACCTTTGGCCCTTTGGCTACAGAAACCCAT	5442
Db	5536	CGCA CAGACGATCTCAAGTGGAGTCGTTTACCTTTGGCCCTTTGGCTACAGAAACCCAT	5595
Qy	5443	CTGTTTGTGACCCACTTCAACTACTCGAAGTAATGAGATCCAGGCACGCTCTCTAGCA	5502
Db	5596	CTGTTTGTGACCCACTTCAACTACTCGAAGTAATGAGATCCAGGCACGCTCTCTAGCA	5655
Qy	5503	GGGAACCCCTGCCCAGAGGTACTTGACATCCCGAACCCGGGCTACCTGGGCCCTGCCATT	5562
Db	5656	GGGAACCCCTGCCCAGAGGTACTTGACATCCCGAACCCGGGCTACCTGGGCCCTGCCATT	5715
Qy	5563	TCCTCAGAGGGGATTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTTGCTGC	5622
Db	5716	TCCTCAGAGGGGATTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTTGCTGC	5775
Qy	5623	AAGGGAAACCTCGTGAAGGAGTCGGGCACCTGAACACACCGGGGCCCGTCCACCTCCCGC	5682
Db	5776	AAGGGAAACCTCGTGAAGGAGTCGGGCACCTGAACACACCGGGGCCCGTCCACCTCCCGC	5835
Qy	5683	AGTAGCCCCAAAGCAGAGGCCACCCACGTTACCAACGAGCACATCAACAGCGGTGGCC	5742
Db	5836	AGTAGCCCCAAAGCAGAGGCCACCCACGTTACCAACGAGCACATCAACAGCGGTGGCC	5895
Qy	5743	TCBAGCCACGCGCCCGGAGGCCACAGCCACCCGCGAGAGCAGACACACCCACCGC	5802
Db	5896	TCBAGCCACGCGCCCGGAGGCCACAGCCACCCGCGAGAGCAGACACACCCACCGC	5955
Qy	5803	TACCGCAGGGGGCAGCCGAGCTGCGCAGGGAACAAGTCTCTGGCCGCCCCCTGGAGCGA	5862
Db	5956	TACCGCAGGGGGCAGCCGAGCTGCGCAGGGAACAAGTCTCTGGCCGCCCCCTGGAGCGA	6015
Qy	5863	GAGNAGTCCCCCGCCGGATGCTCAGCACCGGGAGAGAGCGGTCCCCCGGGAGGCTGTTT	5922
Db	6016	GAGAAGTCCCCCGCCGGATGCTCAGCACCGGGAGAGAGCGGTCCCCCGGGAGGCTGTTT	6075
Qy	5923	GAAGACAGCAGAGGGGCGGCTGCCTGCGGGAGCCGCTGAGGAGCCCCCGCTGTCCCAAGTG	5982
Db	6076	GAAGACAGCAGAGGGGCGGCTGCCTGCGGGAGCCGCTGAGGAGCCCCCGCTGTCCCAAGTG	6135
Qy	5983	AACAAGGGAAGAGGGCAGAGTGCTCTCAAGTTTTCACGGTTTAACATGTCACTATATAT	6042
Db	6136	AACAAGGTCTGGACACAGTCTTCAGTATAAATCTCAGCCAGAAAAACCACTCCTCATCT	6195

RESULT 3

AAL55217

AAL55217
ID AAL55217 standard; DNA; 6156 BP.

[illegible]

AC

AC XX
AAL:

01-MAY-2003 (first entry)

XX
XX
1-70

Human CR1K related DNA sequence. SEQ ID No. 8.

XX
XX
XX

XX Anorectic: hypotensive: cardiant: antilipaemic: cerebroprotective: KW

antiquout; osteopathic; antiarthritic; cytostatic; antidepressant; KW

[illegible]

neuroprotective; antiinflammatory; antidiabetic; analgesic;

human citron rho/rac-interacting kinase; enzyme; CRUK; ameliorating; KW

KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;

KW central nervous system disorder; chronic obstructive pulmonary disease;

KW diabetes; pain; ds.

XX

OS Homo sapiens.

XX

WO2003004523-A1.
 16-JAN-2003.
 28-JUN-2002; 2002WO-EP007156.
 02-JUL-2001; 2001US-0301841P.
 11-DEC-2001; 2001US-0338651P.
 25-APR-2002; 2002US-0375014P.
 (FARB) BAYER AG.
 Zhu Z;
 WPI; 2003-221576/21.
 New human citron rho/rac-interacting kinase (CRIK) polypeptide and polynucleotide, useful in preventing, ameliorating or treating diseases associated with human CRIK dysfunction, e.g. obesity, diabetes or Alzheimer's disease.
 Disclosure; Page 217-222; 237pp; English.
 The invention relates to an isolated polynucleotide encoding a human citron rho/rac-interacting kinase polypeptide. The isolated polynucleotide comprises a 6165 or 8603 base pair sequence, given in the specification. The human citron rho/rac-interacting kinase (CRIK) polypeptide and polynucleotide are useful in preventing, ameliorating, or treating diseases associated with human CRIK dysfunction such as obesity and obesity-associated comorbidities (e.g. hypertension, coronary artery disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of cancer including endometrial, breast, prostate and colon cancer), anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood disorders, anxiety disorders, Parkinson's disease, or diabetes. These can also be used to treat pain associated with the disorders. The human CRIK polypeptide is also useful in diagnostic assays or in genetic testing. The expression vector or the reagent is useful in preparing a medicament for modulating the activity of a human CRIK in a disease, e.g. obesity, a central nervous system disorder, or chronic obstructive pulmonary disease. The fusion protein is useful for generating antibodies against a CRIK polypeptide and for use in various assay systems. The methods are useful in producing and detecting the polynucleotide and polypeptide and in screening for agents that modulate the activity of the human CRIK polypeptide. This polynucleotide sequence represents a DNA sequence relating to the human CRIK protein of the invention
 Sequence 6156 BP: 1732 A: 1548 C: 1679 G: 1197 T: 0 U: 0 Other:

Human CR1K related DNA sequence, SEQ ID No 8.

QY 301 AGTCTTTGAGTTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAAACACCGGG 360
Db 301 AGTCTTTGAGTTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAAACACCGGG 360
QY 361 GACATCTATGCTATGAAAGTGTGAAGAGAGAGAGGCTTTATTTGGCCCGCAGGAGCAGATTTC 420
Db 361 GACATCTATGCTATGAAAGTGTGAAGAGAGAGAGGCTTTATTTGGCCCGCAGGAGCAGATTTC 420
QY 421 TTTTGTGAGAAAGAGCGGACATATTTATCTCAAGCACAAAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGAAAGAGCGGACATATTTATCTCAAGCACAAAGCCCGTGGATCCCCCAATTA 480
QY 481 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
Db 481 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
QY 541 GACTTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600
Db 541 GACTTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTACACGCTTCACTGATGGGATACGTCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTACACGCTTCACTGATGGGATACGTCATCGA 660
QY 661 GACATCAAGCCTGAGAACATTTCTCGTGTACCCGACAGGACACATCAAGCTGGTGAATTT 720
Db 661 GACATCAAGCCTGAGAACATTTCTCGTGTACCCGACAGGACACATCAAGCTGGTGAATTT 720
QY 721 GGATCTGCCCGCGAAATGAATTTCAACAGAGTGGTGAATGCCAACTCCCGATTGGGACC 780
Db 721 GGATCTGCCCGCGAAATGAATTTCAACAGAGTGGTGAATGCCAACTCCCGATTGGGACC 780
QY 781 CCAGATTACATGGCTCTCGAAGTGTGATGAGGAGGATGAAAGGACACCTAC 840
Db 781 CCAGATTACATGGCTCTCGAAGTGTGATGAGGAGGATGAAAGGACACCTAC 840
QY 841 GGCCTGGACTGTGACTGGTGGTCACTGGGCGGTGATTGCCATGAGATGATTATGGGAGA 900
Db 841 GGCCTGGACTGTGACTGGTGGTCACTGGGCGGTGATTGCCATGAGATGATTATGGGAGA 900
QY 901 TCCCGCTTCGCAGAGGGAACCTCTGCAGAACCTTCAATAACATPATGAATTTCCAGCGG 960
Db 901 TCCCGCTTCGCAGAGGGAACCTCTGCAGAACCTTCAATAACATPATGAATTTCCAGCGG 960
QY 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATCAAAAGC 1020
Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATCAAAAGC 1020
QY 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
Db 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
QY 1081 TCTAAAATTGACTGGAAACACATTCGTAACTCTCTCCGCCCTTGGTTCACACCTCAAG 1140
Db 1081 TCTAAAATTGACTGGAAACACATTCGTAACTCTCTCCGCCCTTGGTTCACACCTCAAG 1140
QY 1141 TCTGACGATGACACTCCCAATTTGATGAACACAGAGAGAAATTCGTGGGTTTCATCCCT 1200
Db 1141 TCTGACGATGACACTCCCAATTTGATGAACACAGAGAGAAATTCGTGGGTTTCATCCCT 1200
QY 1201 CCGTGCCAGCTGAGCCCTTCAGGCTTCTGGGTGAAGAACTGCGCTTTGTGGGTTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCTTCAGGCTTCTGGGTGAAGAACTGCGCTTTGTGGGTTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGATTCCTGGTAGATCTGATGCTGTGTGTCGGGTCTGACTCC 1320
Db 1261 TACAGCAAGGCACTGGGATTCCTGGTAGATCTGATGCTGTGTGTCGGGTCTGACTCC 1320
QY 1321 CCTGCCAAGACTAGCTCCATGGAAGAAATTTCTCATCAAAAGCAAAAGAGCTACAAGAC 1380
Db 1321 CCTGCCAAGACTAGCTCCATGGAAGAAATTTCTCATCAAAAGCAAAAGAGCTACAAGAC 1380
QY 1381 TCTCAGGACAAAGTGTACAAAGATGAGCGGAAATGACCCGGTTACATCGGAGAGTGTCA 1440

Db 1381 TCTCAGGACAAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTACATCGAGAGTGTCA 1440
QY 1441 GAGCTGGAGGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1441 GAGCTGGAGGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
QY 1501 TCCCTCTCTGGAGCAGGACCTTGTCTACCTACATCACAAGATGCAGTAGCTTTAAAGCGAAGT 1560
Db 1501 TCCCTCTCTGGAGCAGGACCTTGTCTACCTACATCACAAGATGCAGTAGCTTTAAAGCGAAGT 1560
QY 1561 TTGAGCAAGCAGCAGGATGGAGGTGTCCAGGAGGATCACAAGCACTGCGAGCTTCTCCAT 1620
Db 1561 TTGAGCAAGCAGCAGGATGGAGGTGTCCAGGAGGATCACAAGCACTGCGAGCTTCTCCAT 1620
QY 1621 GATATCAGAGCAGCAGGACCGGAGGCTCCAAAGAAATCAAAGCAGGAGTAGTACCAGGCTCAA 1680
Db 1621 GATATCAGAGCAGCAGGACCGGAGGCTCCAAAGAAATCAAAGCAGGAGTAGTACCAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGTTGATGATGAATCACTTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740
Db 1681 GTGGAAGAAATGAGTTGATGATGAATCACTTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740
QY 1741 CGGAGTCACTCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGATTTCAAG 1800
Db 1741 CGGAGTCACTCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGATTTCAAG 1800
QY 1801 CGGAAAGCGCAGAGATGTCAAGTAAACTTGTGAAGGCTAAGGATCAAGGAGGACCTGAA 1860
Db 1801 CGGAAAGCGCAGAGATGTCAAGTAAACTTGTGAAGGCTAAGGATCAAGGAGGACCTGAA 1860
QY 1861 GTGGGAAATATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
Db 1861 GTGGGAAATATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
QY 1921 CTCCAAAGAGAAACTTGGAGAAAGGCT----- 1944
Db 1921 CTCCAAAGAGAAACTTGGAGAAAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAG 1980
QY 1945 -----GCAAGAGAGCGAGCGAGAGGAGCTGGAGAAAGCTTGAGAACCGAGAG 1992
Db 1981 AATATCGCGCAGGCAAGAGGAGCGAGCGAGGAGGAGCTGGAGAAAGCTTGAGAACCGAGAG 2040
QY 1993 GATTCTTCTGAAGCATCAGAAAGAGAGCTGCTGGAAGCTGAGGAACGCGCCCAATTTCTCTG 2052
Db 2041 GATTCTTCTGAAGCATCAGAAAGAGAGCTGCTGGAAGCTGAGGAACGCGCCCAATTTCTCTG 2100
QY 2053 GAGAACAAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGGATGAC 2112
Db 2101 GAGAACAAAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGGATGAC 2160
QY 2113 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGTGATAAAATTTCTGGAGCTGAGAG 2172
Db 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGTGATAAAATTTCTGGAGCTGAGAG 2220
QY 2173 AAACATCGGAGGCGCCCAAGTCTCAGCCCAAGCTAGAAAGTGCACCTGAAAACAGAAAGAG 2232
Db 2221 AAACATCGGAGGCGCCCAAGTCTCAGCCCAAGCTAGAAAGTGCACCTGAAAACAGAAAGAG 2280
QY 2233 CAGCACTATCAGGAAAGAGATTAAAGTTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2292
Db 2281 CAGCACTATCAGGAAAGAGATTAAAGTTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340
QY 2293 AAGGAGACCTGAGAAACATGATGAGAGACAGAGGAGGAGGCCCATGAGAAAGGCAAA 2352
Db 2341 AAGGAGACCTGAGAAACATGATGAGAGACAGAGGAGGAGGCCCATGAGAAAGGCAAA 2400
QY 2353 ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2412
Db 2401 ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2460
QY 2413 CAGAGGATTTGGAACTGTCTGAGGCCAATAAATTTGACAGCAATAGCAGTCTTTTACC 2472

QY 4633 GCA----- 4635
DB 4636 GCAGATGTCCCATATACATCTGAAGATGGAATCTCACCCGCACACCACTGCTGGCCCGGG 4695
QY 4636 ----- 4635
DB 4696 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAAAGCGCTGGGTCAACGCCCTTA 4755
QY 4636 -----GAAAAAGCAGAAAGCTGATCTAAACTG 4662
DB 4756 GAATCAGTTGTGCGAGGTGGAGAGTCTTCTAGGGNAAGAGCAGAAAGCTGATCTAAACTG 4815
QY 4663 CTTGGAAACTCCCTGCTGAACTGGAAGGTGATGACCGCTGAGACATGAACCTGACACCTG 4722
DB 4816 CTTGGAAACTCCCTGCTGAACTGGAAGGTGATGACCGCTGAGACATGAACCTGACACCTG 4875
QY 4723 CTTTCACTGACCAAGTGGTGTGTGTGGGCAACCGAGGAAGGGCTCTACGCCCTGAATGTC 4782
DB 4876 CTTTCACTGACCAAGTGGTGTGTGTGGGCAACCGAGGAAGGGCTCTACGCCCTGAATGTC 4935
QY 4783 TTGAAAACTCCCTAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4842
DB 4936 TTGAAAACTCCCTAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995
QY 4843 AAGGACCTGGAGAACTACTCATGATAGCAGAGAGAGCGGGCACTGTCTTTGTGGAC 4902
DB 4996 AAGGACCTGGAGAACTACTCATGATAGCAGAGAGAGCGGGCACTGTCTTTGTGGAC 5055
QY 4903 GTGAAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTTGTCTGCTGCCAGCCGACATCTCA 4962
DB 5056 GTGAAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTTGTCTGCTGCCAGCCGACATCTCA 5115
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTTGGGGCAGGCAAGATTGAGAAC 5022
DB 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTTGGGGCAGGCAAGATTGAGAAC 5175
QY 5023 GGGCTCTGCATCTGTGACGCCATGCCAGCAAGTCTGTCATTCTCCGCTACAAAGAAAC 5082
DB 5176 GGGCTCTGCATCTGTGACGCCATGCCAGCAAGTCTGTCATTCTCCGCTACAAAGAAAC 5235
QY 5083 CTCAGCAAACTAGCATCCGGAAAGATAGAGACTCGAGCCCTGCGAGCTGTATCCAC 5142
DB 5236 CTCAGCAAACTAGCATCCGGAAAGATAGAGACTCGAGCCCTGCGAGCTGTATCCAC 5295
QY 5143 TTCACCAATTACAGTATCTCTATTGAAACCAATAAATCTACGAAATCGACATGAAGCAG 5202
DB 5296 TTCACCAATTACAGTATCTCTATTGAAACCAATAAATCTACGAAATCGACATGAAGCAG 5355
QY 5203 TACAGCTCGAGGAATTCCTGGATAGAATGACCAATTCCTTGGCACTGCTGTGTTGCC 5262
DB 5356 TACAGCTCGAGGAATTCCTGGATAGAATGACCAATTCCTTGGCACTGCTGTGTTGCC 5415
QY 5263 GCCTCTCCAAAGCTTCCCTGCTCTCAATCGTGCAGGTGCAAGCGGCGAGCGAGAG 5322
DB 5416 GCCTCTCCAAAGCTTCCCTGCTCTCAATCGTGCAGGTGCAAGCGGCGAGCGAGAG 5475
QY 5323 GAGTACTTGTGTTTCCACGAATTTGGAGTGTTCGTGGAATTTTACGGAAAGAGCTAGC 5382
DB 5476 GAGTACTTGTGTTTCCACGAATTTGGAGTGTTCGTGGAATTTTACGGAAAGAGCTAGC 5535
QY 5383 CGCACAGACATCTCAAGTGGAGTCTGCTTACCTTTTGGCCCTTTGCTTACAGAGAACCTAT 5442
DB 5536 CGCACAGACATCTCAAGTGGAGTCTGCTTACCTTTTGGCCCTTTGCTTACAGAGAACCTAT 5595
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCAAGCTCTCCAGCA 5502
DB 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCAAGCTCTCCAGCA 5655
QY 5503 GGGACCCCTGCGGAGGCTACTTGGACATCCCGAAACCCGGGCTACTGGGCGCTGCAATT 5562
DB 5656 GGGACCCCTGCGGAGGCTACTTGGACATCCCGAAACCCGGGCTACTGGGCGCTGCAATT 5715
QY 5563 TCCTCAGGAGGATTTACTTGGCGTCTCATACAGGATAAATTAAGGGTCAATTGCTGC 5622

DB 5716 TCTCAGGAGCGATTACTTGGCTCTCATACAGGATAAATTAAGGGTCAATTGCTGC 5775
QY 5623 AAGGAAAACCTCGTGAAGAGTCCCGCACTGAACCAACCGGGCGCGTCCACCTCCCGC 5682
DB 5776 AAGGAAAACCTCGTGAAGAGTCCCGCACTGAACCAACCGGGCGCGTCCACCTCCCGC 5835
QY 5683 AGCAGCCCCAAAGCGAGGCCCAACCCAGTACACGAGCACATCAACAAAGCGGTGGCC 5742
DB 5836 AGCAGCCCCAAAGCGAGGCCCAACCCAGTACACGAGCACATCAACAAAGCGGTGGCC 5895
QY 5743 TCAGCCCCAGCGCCCGCGAAGGCCACCGACCCCGAGAGCCAGCACACCCACCGC 5802
DB 5896 TCAGCCCCAGCGCCCGCGAAGGCCACCGACCCCGAGAGCCAGCACACCCACCGC 5955
QY 5803 TACCGGAGGGGGAGCCAGAGTGGCAGGACCAAGTCTCTGGCGCGCCCTTGGAGCGA 5862
DB 5956 TACCGGAGGGGGAGCCAGAGTGGCAGGACCAAGTCTCTGGCGCGCCCTTGGAGCGA 6015
QY 5863 GAGAAGTCCCGCGCGGATGTCTACAGCGGAGAGAGGGTCTCCCGGGAGGCTGTTT 5922
DB 6016 GAGAAGTCCCGCGCGGATGTCTACAGCGGAGAGAGGGTCTCCCGGGAGGCTGTTT 6075
QY 5923 GAAGACAGCAGCAGGGCGCGCTGCTGGGAGCGGTGAGGACCCCGCTGTCCACAGTG 5982
DB 6076 GAAGACAGCAGCAGGGCGCGCTGCTGGGAGCGGTGAGGACCCCGCTGTCCACAGTG 6135
QY 5983 AACAAAGGAAGAGGGCA 5999
DB 6136 AACAAAGTGAAGCGCA 6152
RESULT 4
AAD38864
ID AAD38864 standard; cDNA; 6298 BP.
XX AAD38864;
AC AAD38864;
XX 23-SEP-2002 (first entry)
XX Human kinase (PKIN)-21 cDNA.
XX Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
KW development; hepatitis; cardiovascular; hypertension; drug screening;
KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
KW hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-hiv;
KW neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
KW hyperlipidaemia; enzyme; gene; 55.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 55..6219
XX /*tag= a
XX /product= "Human kinase (PKIN)-21"
XX
XX PN WO200233099-A2.
XX
XX PD 25-APR-2002.
XX
XX XX 20-OCT-2001; 2001WO-US047728.
XX
XX XX 20-OCT-2000; 2000US-0242410P.
XX PR 27-OCT-2000; 2000US-0244068P.
XX PR 03-NOV-2000; 2000US-0245708P.
XX PR 09-NOV-2000; 2000US-0247672P.
XX PR 16-NOV-2000; 2000US-0249565P.
XX PR 22-NOV-2000; 2000US-0252730P.
XX PR 01-DEC-2000; 2000US-0250807P.
XX

Db 1555 TCCCTCTCGGACGAGCCTTGCTACCTACATCAGAAATGCAGTAGCTTAAAGCGAGT 1614
Qy 1561 TTGGAGCAAGCACGATGGAGTGTCCAGAGGATGA CAAAGCATCTCAGCTTCTCCAT 1620
Db 1615 TTGGAGCAAGCACGATGGAGTGTCCAGAGGATGA CAAAGCATCTCAGCTTCTCCAT 1674
Qy 1621 GATATCAGAGACAGACGGAGCTCCAAGAAATCAAAGAGCAGAGGTACAGGCTCAA 1680
Db 1675 GATATCAGAGACAGACGGAGCTCCAAGAAATCAAAGAGCAGAGGTACAGGCTCAA 1734
Qy 1681 GTGGAAGAAATGAGTGTGATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740
Db 1735 GTGGAAGAAATGAGTGTGATGATGAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1794
Qy 1741 CGGAGTGATCTTACGAATCTGAGCTGAGAGGTCTCGGCTTGCTGCTGAAGAAATTCAG 1800
Db 1795 CGGAGTGATCTTACGAATCTGAGCTGAGAGGTCTCGGCTTGCTGCTGAAGAAATTCAG 1854
Qy 1801 CGGAAAGCGACAGAAATGTGACATAAATCTGTTGAAGGCTTAAGGATCAAGGGAAGCCTGAA 1860
Db 1855 CGGAAAGCGACAGAAATGTGACATAAATCTGTTGAAGGCTTAAGGATCAAGGGAAGCCTGAA 1914
Qy 1861 GTGGAGAAATATGCGAAATCTGAGAAAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1915 GTGGAGAAATATGCGAAATCTGAGAAAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGAG 1974
Qy 1921 CTCCAAGAGAACTGGAGAGGCTG----- 1945
Db 1975 CTCCAAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCCAGGCTGCTGCAG 2034
Qy 1946 -----CAAAGGAGCGAGCCGAGAGGGAGCTGGAGAGCTCGAAGCTCGAAGCCGAGAG 1992
Db 2035 AATATCCGCGCAGGCAAAAGGAGCGAGCGAGAGGGAGCTGGAGAGCTCGAAGCCGAGAG 2094
Qy 1993 GATTTCTTGAAGGCATCAGAAAGAGCTGTGGAGCTGAGAAAGCGGCGCATTTCTGT 2052
Db 2095 GATTTCTTGAAGGCATCAGAAAGAGCTGTGGAGCTGAGAAAGCGGCGCATTTCTGT 2154
Qy 2053 GAGAACAGGTAAGAGACTAGAGACCTAGAGCCATGGAGCGTAGAGAAACAGAGCTGAAGGATGAC 2112
Db 2155 GAGAACAGGTAAGAGACTAGAGACCTAGAGCCATGGAGCGTAGAGAAACAGAGCTGAAGGATGAC 2214
Qy 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2172
Db 2215 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2274
Qy 2173 AAACATCGGGAGGCCCAAGTCTCAGCCAGCAGCCTAGAGTGCACCTGAAACAGAAAGAG 2232
Db 2275 AAACATCGGGAGGCCCAAGTCTCAGCCAGCAGCCTAGAGTGCACCTGAAACAGAAAGAG 2334
Qy 2233 CAGCACTATCAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGCTGAC 2292
Db 2335 CAGCACTATCAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGCTGAC 2394
Qy 2293 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGAGGAGGCCCATCAGAGAGGCCAA 2352
Db 2395 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGAGGAGGCCCATCAGAGAGGCCAA 2454
Qy 2353 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATCGATTCCAAGATCAGATCCCTGGAA 2412
Db 2455 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATCGATTCCAAGATCAGATCCCTGGAA 2514
Qy 2413 CAGAGGATTTGTGAACTGTCTGAAGCCATAAATCTTGACGCAATAGCAGTCTTTTACC 2472
Db 2515 CAGAGGATTTGTGAACTGTCTGAAGCCATAAATCTTGACGCAATAGCAGTCTTTTACC 2574
Qy 2473 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
Db 2575 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2634
Qy 2533 CTGGAGACACAGGCTGGAGTTGGAGGCCAGAACCGGAAATCTGGAGAGCAGCTGGAG 2592
Db 2635 CTGGAGACACAGGCTGGAGTTGGAGGCCAGAACCGGAAATCTGGAGAGCAGCTGGAG 2694

Qy 2593 AAGATCAGCCACCAAGACCACTGACAGAATCGCTGCTGGAATCGGAGACAAGATTG 2652
Db 2695 AAGATCAGCCACCAAGACCACTGACAGAATCGCTGCTGGAATCGGAGACAAGATTG 2754
Qy 2653 CGGAGGTCTAGTCTTAGAGCACGAGGAGCGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2712
Db 2755 CGGAGGTCTAGTCTTAGAGCACGAGGAGCGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2814
Qy 2713 CTACAGCTCTCCCTGCAAGGAGCGGAGTCAAGTTGACAGCCTGCAAGGCTGCAAGCGCG 2772
Db 2815 CTACAGCTCTCCCTGCAAGGAGCGGAGTCAAGTTGACAGCCTGCAAGGCTGCAAGCGCG 2874
Qy 2773 GCCCTGAGAGCGCAGCTTCCGACGAGGAGACAGAGCTGGAAGAGACCAAGCT 2832
Db 2875 GCCCTGAGAGCGCAGCTTCCGACGAGGAGACAGAGCTGGAAGAGACCAAGCT 2934
Qy 2833 GAAGAGAGATTCAGGCACTCAGCGCACATAGAGATGAATAATCCAGCGCAAAATTTGATCT 2892
Db 2935 GAAGAGAGATTCAGGCACTCAGCGCACATAGAGATGAATAATCCAGCGCAAAATTTGATCT 2994
Qy 2893 CTTTCGTAACAGCTGTACTGTAATCAGAGACCTCGAGGAGCAGCTTAAACCAAGCTGACCGAG 2952
Db 2995 CTTTCGTAACAGCTGTACTGTAATCAGAGACCTCGAGGAGCAGCTTAAACCAAGCTGACCGAG 3054
Qy 2953 GACAAACGCTGAACCTCAAACCAAACTTCTACTTTGTCCTGTCCTGATGAGGCTTCT 3012
Db 3055 GACAAACGCTGAACCTCAAACCAAACTTCTACTTTGTCCTGTCCTGATGAGGCTTCT 3114
Qy 3013 GGCGCCAAACAGAGATTTGTAACCTGCGAAGTGAAGTGAAGCCATCTCCGCCGGGAGATC 3072
Db 3115 GGCGCCAAACAGAGATTTGTAACCTGCGAAGTGAAGTGAAGCCATCTCCGCCGGGAGATC 3174
Qy 3073 ACGGAAACAGAGATTCAGCTTACCGAGCCAGAAACAGATCGAGGCTCTGAAGACCAAG 3132
Db 3175 ACGGAAACAGAGATTCAGCTTACCGAGCCAGAAACAGATCGAGGCTCTGAAGACCAAG 3234
Qy 3133 TGCAACCATGCTGGAGGAAACAGGTCATGGATTGGAGGCCCTTAAACGATGAGCTGTAGAA 3192
Db 3235 TGCAACCATGCTGGAGGAAACAGGTCATGGATTGGAGGCCCTTAAACGATGAGCTGTAGAA 3294
Qy 3193 AAACAGCGGAGTGGGAGGCTCGAGGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGAG 3252
Db 3295 AAACAGCGGAGTGGGAGGCTCGAGGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGAG 3354
Qy 3253 TGTCCGGTTCCAGAGCTGACAGAAATCTCGAACACCGAGAAACAGAGCGGCGAGAGCC 3312
Db 3355 TGTCCGGTTCCAGAGCTGACAGAAATCTCGAACACCGAGAAACAGAGCGGCGAGAGCC 3414
Qy 3313 GATCAGCGGATCACCGAGTCTCCCGAGTGGAGCTGAGCTGAGAGGAGCAAGGCT 3372
Db 3415 GATCAGCGGATCACCGAGTCTCCCGAGTGGAGCTGAGCTGAGAGGAGCAAGGCT 3474
Qy 3373 GAGATTCTCGCTCTGACGAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3432
Db 3475 GAGATTCTCGCTCTGACGAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3534
Qy 3433 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGCTTA 3492
Db 3535 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGCTTA 3594
Qy 3493 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGACCAA 3552
Db 3595 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGACCAA 3654
Qy 3553 TTAACAGCAGAGATGGAACCTGCAAGAAATCAACATTTCCGCTGCTCAAGGAGCTGAA 3612
Db 3655 TTAACAGCAGAGATGGAACCTGCAAGAAATCAACATTTCCGCTGCTCAAGGAGCTGAA 3714
Qy 3613 GAACTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTACTTGGAGTATCAGCTG 3672
Db 3715 GAACTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTACTTGGAGTATCAGCTG 3774

QY 3673 GAAACATTTCAGGTTCTCTATTCTCATGAAAGGTTGAATGGAAGCGCATTTTCTCAA 3732
Db 3775 GAAACATTTCAGGTTCTCTATTCTCATGAAAGGTTGAATGGAAGCGCATTTTCTCAA 3834
QY 3733 CAACCAAACTCATGTTTTCTGCAAGCCAAATGACCAACTCTGCTAAAGAAAG 3792
Db 3835 CAACCAAACTCATGTTTTCTGCAAGCCAAATGACCAACTCTGCTAAAGAAAG 3892
QY 3793 GGTTTATTAGTCGACGGAAGAGGCCCTGCTTTACCCACACAGGTTCTCTGCAATAC 3852
Db 3893 -----AGTTTCTCTGCAATAC 3909
QY 3853 AATGAGCTGAAGCTGGCCCTGGAGAAGGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3912
Db 3910 AATGAGCTGAAGCTGGCCCTGGAGAAGGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3969
QY 3913 CTTCAAGAACCCGATCGAGTCCGGTCCGCCGGGAGGAAGCTGCCACCGCAAGCA 3972
Db 3970 CTTCAAGAACCCGATCGAGTCCGGTCCGCCGGGAGGAAGCTGCCACCGCAAGCA 4029
QY 3973 ACGGACACCCACACCATCCACGCGAGCCACGCGAGGAGCAGATCGCCATGTCGCGC 4032
Db 4030 ACGGACACCCACACCATCCACGCGAGCCACGCGAGGAGCAGATCGCCATGTCGCGC 4089
QY 4033 ATCGTGGGTGCGCAGACACGACCCAGTGCATGAGCTGCTGGCCCGCCATCCAGC 4092
Db 4090 ATCGTGGGTGCGCAGACACGACCCAGTGCATGAGCTGCTGGCCCGCCATCCAGC 4149
QY 4093 CGCAGAAAGAGTCTCACTCCAGAGGAATTTAGTCCGGCTTTAAGGAAAGCATGAC 4152
Db 4150 CGCAGAAAGAGTCTCACTCCAGAGGAATTTAGTCCGGCTTTAAGGAAAGCATGAC 4209
QY 4153 CACAATATTCTCACCAGATTCAAAGTGGAGTGAACATCGAGGCCACAAAGTGTGTG 4212
Db 4210 CACAATATTCTCACCAGATTCAAAGTGGAGTGAACATCGAGGCCACAAAGTGTGTG 4269
QY 4213 TGTCTGGATACGTCGACATTGGACCCAGGATCCAAATGCTCGAATGTGAGTGATG 4272
Db 4270 TGTCTGGATACGTCGACATTGGACCCAGGATCCAAATGCTCGAATGTGAGTGATG 4329
QY 4273 TGTCAACCCAAAGTCTCCAGCTGCTTGCAGCCACCTGCGGCTTGCTGCTCAATATG 4332
Db 4330 TGTCAACCCAAAGTCTCCAGCTGCTTGCAGCCACCTGCGGCTTGCTGCTCAATATG 4389
QY 4333 ACACATTCAACGAGGCTTCTGCGGTGACAAATGAACCTCCCAAGTCTCAGACCAAG 4392
Db 4390 ACACATTCAACGAGGCTTCTGCGGTGACAAATGAACCTCCCAAGTCTCAGACCAAG 4449
QY 4393 GAGCCAGCAGCAGCTGACCTTGAAGGAGTGAAGAGTCCAGGAATACCAAGCA 4452
Db 4450 GAGCCAGCAGCAGCTGACCTTGAAGGAGTGAAGAGTCCAGGAATACCAAGCA 4509
QY 4453 GGAACAGAGGCTGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCTATTAT 4512
Db 4510 GGAACAGAGGCTGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCTATTAT 4569
QY 4513 GACAAAGAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTCCTCC 4572
Db 4570 GACAAAGAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTCCTCC 4629
QY 4573 GAGGGGATGATCTATTATGCTGCTGCTGCTTCCGAATCGCAAAATACAGCCAAA 4632
Db 4630 GAGGGGATGATCTATTATGCTGCTGCTGCTTCCGAATCGCAAAATACAGCCAAA 4689
QY 4633 GCA----- 4635
Db 4690 GCAGATGTCCCATACATACCTGAAGTGAATCTCACCCGACACACCTGCTGGCCGG 4749
QY 4636 ----- 4635
Db 4750 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4809
QY 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662

Db 4810 GAATCAGTTGTCGAGGTGGGAGAGTTTCTAGGGAAGAAAGCAGCTGATGCTAAACTG 4859
QY 4663 CTTGAAACTCCCTGCTGCTGAAACTGAAAGGTGATGACCGTCTTAGACATGAACTGCACGCTG 4722
Db 4870 CTTGAAACTCCCTGCTGAAACTGAAAGGTGATGACCGTCTTAGACATGAACTGCACGCTG 4929
QY 4723 CCCTTCAGTGAACGAGTGGTGTGGTGGGCACAGAGGAGGGCTCTAGCCCTGAATGTC 4782
Db 4930 CCCTTCAGTGAACGAGTGGTGTGGTGGGCACAGAGGAGGGCTCTAGCCCTGAATGTC 4989
QY 4783 TTGAAAACTCCCTAAACCCATGTCCAGGAATTCGAGCAGTCTTCCAAATTTATATTATC 4842
Db 4990 TTGAAAACTCCCTAAACCCATGTCCAGGAATTCGAGCAGTCTTCCAAATTTATATTATC 5049
QY 4843 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAAAGGCGGACATGTGTCTTGTGAC 4902
Db 5050 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAAAGGCGGCACTGTGTCTTGTGAC 5109
QY 4903 GTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4962
Db 5110 GTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5169
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGTGCACATTTGTTGGGCGAGCAAGATTGAAGAC 5022
Db 5170 CCCAACATTTTGAAGCTGTCAAGGGTGCACATTTGTTGGGCGAGCAAGATTGAAGAC 5229
QY 5023 GGGCTCTGCACTCTGTGAGCAGCATGCCAGCAAGTCTGCTCATTTCTCGCTACACGAAAC 5082
Db 5230 GGGCTCTGCACTCTGTGAGCAGCATGCCAGCAAGTCTGCTCATTTCTCGCTACACGAAAC 5289
QY 5083 CTCAGCAATACTGCATCCCGAAAGAGATAGACCTTCAGAGCCCTGACGTGTATCCAC 5142
Db 5290 CTCAGCAATACTGCATCCCGAAAGAGATAGACCTTCAGAGCCCTGACGTGTATCCAC 5349
QY 5143 TTCAACCAATTACAGTATCTCTATGGNACCAATAAATTTCTACGAATTCGACATGAGCAG 5202
Db 5350 TTCAACCAATTACAGTATCTCTATGGNACCAATAAATTTCTACGAATTCGACATGAGCAG 5409
QY 5203 TACACGCTCCAGGAATTTCTGGATAGCAATTTCTTGGCAGCTGCTGTGTTGCC 5262
Db 5410 TACACGCTCCAGGAATTTCTGGATAGCAATTTCTTGGCAGCTGCTGTGTTGCC 5469
QY 5263 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGAGGCGAGAG 5322
Db 5470 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGAGGCGAGAG 5529
QY 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTGTTGATTTCTTACGGAAGACGTAGC 5382
Db 5530 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTGTTGATTTCTTACGGAAGACGTAGC 5589
QY 5383 CGCACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTCCCTACAGAGAACCTAT 5442
Db 5590 CGCACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTCCCTACAGAGAACCTAT 5649
QY 5443 CTGTTTGTGACCCACCTCACTCACTCGAGTAAATTTGAGATCCAGGACGCTCCTCAGCA 5502
Db 5650 CTGTTTGTGACCCACCTCACTCACTCGAGTAAATTTGAGATCCAGGACGCTCCTCAGCA 5709
QY 5503 GGGACCCCTGCCAGCGGTACCTGGACATCCCGAACCCGCGCTTACCTGGGCGCTTGCAT 5562
Db 5710 GGGACCCCTGCCAGCGGTACCTGGACATCCCGAACCCGCGCTTACCTGGGCGCTTGCAT 5769
QY 5563 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGAGATAAATTAAGGTCTATTGCTGC 5622
Db 5770 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGAGATAAATTAAGGTCTATTGCTGC 5829
QY 5623 AAGGGAACCTCGTGAAGGAGTCCGCGACTGAAACACACCGGGGCGCTCCACCTCCGC 5682
Db 5830 AAGGGAACCTCGTGAAGGAGTCCGCGACTGAAACACACCGGGGCGCTCCACCTCCGC 5889
QY 5683 AGCAGCCCAACAGAGGAGGCCCAACCAAGTACACAGCAGCATACCAAGGCGCTGGCC 5742


```
Db 5890 AGCAGCCCAACAAGCGAGGCCACCACGCTACACGAGCACAATCACCAAGCGGTGGCC 5949
Qy 5743 TCCAGCCAGCGCCGCCGAAGCGCCCGACGACCGCGAGAGCCAGCACACCCACCGC 5802
Db 5950 TCCAGCCAGCGCCGCCGAAGCGCCCGACGACCGCGAGAGCCAGCACACCCACCGC 6009
Qy 5803 TACCGCGAGCGCGACCGAGCTGGCGAGGACAAAGTCTTCCTGGCGCCCGCTGGAGCGA 5862
Db 6010 TACCGCGAGCGCGACCGAGCTGGCGAGGACAAAGTCTTCCTGGCGCCCGCTGGAGCGA 6069
Qy 5863 GAGAAGTCCCGCGCGGATGCTCAGCAGCGCGAGAGCGGTCCCGCGAGGCTGT 5922
Db 6070 GAGAAGTCCCGCGCGGATGCTCAGCAGCGCGAGAGCGGTCCCGCGAGGCTGT 6129
Qy 5923 GAAGCAGCAGCAGCGCGCGGTGCTGCGGAGCGGTGAGCAGCCCGGTGTC 5982
Db 6130 GAAGCAGCAGCAGCGCGCGGTGCTGCGGAGCGGTGAGCAGCCCGGTGTC 6189
Qy 5983 AACAGGGAGGCGAGAGTGCCTCTCAAGTTTCAACGTTTAAACACTGTCACTATTAT 6042
Db 6190 AACAGGTCTGGACCAAGTCTTCAGTATAAATCTCAGCCAGAGAAAAACCAACTCTCTCT 6249

RESULT 5
AAL55214
ID AAL55214 standard; DNA; 6165 BP.
XX
AC AAL55214;
XX
DT 01-MAY-2003 (first entry)
XX
DE Human CR1K encoding DNA sequence, SEQ ID No 1.
XX
KW Anorectic; hypotensive; cardiant; antilipaeamic; cerebroprotective;
KW antitout; osteopathic; antiarthritic; cytostatic; antidepressant;
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;
KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;
KW central nervous system disorder; chronic obstructive pulmonary disease;
KW diabetes; pain; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..6165
FT FT /*tag= a
FT FT /product= "Human CR1K protein"
XX
PN WO2003004523-A1.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002WO-BP007156.
XX
PR 02-JUL-2001; 2001US-0301841P.
PR 11-DEC-2001; 2001US-033851P.
PR 25-APR-2002; 2002US-0375014P.
XX
PA (FARB ) BAYER AG.
XX
PI Zhu Z;
XX
PI WPI; 2003-221576/21.
DR P-PSDB; AAO26959.
DR
XX
XX New human citron rho/rac-interacting kinase (CR1K) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX
XX Example 1; Fig 1; 237pp; English.
XX
XX
```

```
CC The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 5603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CR1K)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CR1K dysfunction, such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
CC cancer including endometrial, breast, prostate and colon cancer),
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC disease), chronic obstructive pulmonary disease, or diabetes. These can
CC also be used to treat pain associated with the disorders. The human CR1K
CC polypeptide is also useful in diagnostic assays or in genetic testing.
CC The expression vector or the reagent is useful in preparing a medicament
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a
CC central nervous system disorder, or chronic obstructive pulmonary
CC disease. The fusion protein is useful for generating antibodies against a
CC CR1K polypeptide and for use in various assay systems. The methods are
CC useful in producing and detecting the polynucleotide and polypeptide and
CC in screening for agents that modulate the activity of the human CR1K
CC polypeptide. This polynucleotide sequence represents a DNA sequence
CC encoding a human CR1K protein of the invention
XX
```

Sequence 6165 BP; 1735 A; 1549 C; 1680 G; 1201 T; 0 U; 0 Other;

Query Match 92.0%; Score 5666.2; DB 8; Length 6165;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 5947; Conservative 0; Mismatches 13; Indels 243; Gaps 3;

```
Qy 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCCTTTGGATCGTGGTGTGACCCCAT 60
Db 1 ATGTTGAAGTTCAAATATGGAGCGCGGAATCCCTTTGGATCGTGGTGTGACCCCAT 60
Qy 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
Qy 121 CAGCAGATGTCCTCTTTCCGAGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 121 CAGCAGATGTCCTCTTTCCGAGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Qy 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTTGTCGGGAAGTAT 240
Db 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTTGTCGGGAAGTAT 240
Qy 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGAAAGGACTTCGAAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGAAAGGACTTCGAAAGTCAGA 300
Qy 301 AGTCTTGTAGTTGTGTCACCTTTGCTGAAGTGCAGGTGTTAGAGAGAAGCAACCGGG 360
Db 301 AGTCTTGTAGTTGTGTCACCTTTGCTGAAGTGCAGGTGTTAGAGAGAAGCAACCGGG 360
Qy 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 420
Qy 421 TTTTTCAGGAGAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCAATTA 480
Db 421 TTTTTCAGGAGAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCAATTA 480
Qy 481 CAGTATGCTTTCCAGGACAAAATCACCCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
Db 481 CAGTATGCTTTCCAGGACAAAATCACCCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
Qy 541 GACTTGTCTGCTTTTGAATAGATATGAGGACAGTTAGATGAAAACCTGTACAGTTT 600
Db 541 GACTTGTCTGCTTTTGAATAGATATGAGGACAGTTAGATGAAAACCTGTACAGTTT 600
Qy 601 TACCTAGCTAGCTGATTTGGCTGTTTCAGAGCTTCATCTGATGGATACGTCATCGA 660
Db 601 TACCTAGCTAGCTGATTTGGCTGTTTCAGAGCTTCATCTGATGGATACGTCATCGA 660
```

QY 661 GACATCAAGCCTGAGAACATTCTCGTTGACCGCACAGGACACATCAAGCTGTGGATTTT 720
DB 661 GACATCAAGCCTGAGAACATTCTCGTTGACCGCACAGGACACATCAAGCTGTGGATTTT 720
QY 721 GGATCTCCCGGAAATGAAATTCACACAGATGGTGAATGCCAAATCCCGATTTGGACC 780
DB 721 GGATCTCCCGGAAATGAAATTCACACAGATGGTGAATGCCAAATCCCGATTTGGACC 780
QY 781 CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACCGGGATGGAAGGCACTTAC 840
DB 781 CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACCGGGATGGAAGGCACTTAC 840
QY 841 GGCTGACATGTGACTGGTGGTCACTGGGGGTGATTTGCCATATGAGATGATTTATGGGAGA 900
DB 841 GGCTGACATGTGACTGGTGGTCACTGGGGGTGATTTGCCATATGAGATGATTTATGGGAGA 900
QY 901 TCCCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGG 960
DB 901 TCCCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGG 960
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTTCAAAGC 1020
DB 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTTCAAAGC 1020
QY 1021 TTGTTGTGGCGCCAGAAAGAGACTGAAGTTTGAAGGCTTTTGCTGCCATCTTTCTTC 1080
DB 1021 TTGTTGTGGCGCCAGAAAGAGACTGAAGTTTGAAGGCTTTTGCTGCCATCTTTCTTC 1080
QY 1081 TCTAAATTTGACTGGAAACAATTCTGTAATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1140
DB 1081 TCTAAATTTGACTGGAAACAATTCTGTAATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1140
QY 1141 TCTGACATGACACCTTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGTTCATCTCT 1200
DB 1141 TCTGACATGACACCTTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGTTCATCTCT 1200
QY 1201 CCGTGCAGCTGAGCCCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGCCCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGGTTTTCG 1260
QY 1261 TAGCAGAGCACTGGGATCTTTGTGATGATCTGAGTCTGTTGTGTCGGGCTTGACTCC 1320
DB 1261 TAGCAGAGCACTGGGATCTTTGTGATGATCTGAGTCTGTTGTGTCGGGCTTGACTCC 1320
QY 1321 CCGTGCAGACTAGCTTCCATGGAAAAAATCTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
DB 1321 CCGTGCAGACTAGCTTCCATGGAAAAAATCTTCTCATCAAAAGCAAGAGCTACAAGAC 1380
QY 1381 TCTCAGGACAAAGTGTCAACAAGTGGAGCGGAAATGACCCGGTTACATCGGAGAGTCA 1440
DB 1381 TCTCAGGACAAAGTGTCAACAAGTGGAGCGGAAATGACCCGGTTACATCGGAGAGTCA 1440
QY 1441 GAGGTGGAGCTGTGCTTAGTCAGAGGAGGTGAGCTCAAGGCTCTGAGCTCAGAGA 1500
DB 1441 GAGGTGGAGCTGTGCTTAGTCAGAGGAGGTGAGCTCAAGGCTCTGAGCTCAGAGA 1500
QY 1501 TCGCTCCTGGAGCAGGACCTTGCTACTACATCACAGAAATGAGTAGCTTTAAAGCGAAGT 1560
DB 1501 TCGCTCCTGGAGCAGGACCTTGCTACTACATCACAGAAATGAGTAGCTTTAAAGCGAAGT 1560
QY 1561 TTGAGCAAGCAGGATGAGGTGTCAGGAGGATGACAAAGCACTGAGCTTCCAT 1620
DB 1561 TTGAGCAAGCAGGATGAGGTGTCAGGAGGATGACAAAGCACTGAGCTTCCAT 1620
QY 1621 GATATCAGAGAGCAGAGCGGAGCTTCCAAAGAAATCAAAAGACAGAGTACCAGGCTCAA 1680
DB 1621 GATATCAGAGAGCAGAGCGGAGCTTCCAAAGAAATCAAAAGACAGAGTACCAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 GTGGAAGAAATGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740

QY 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGATCTCGGCTTGTCTGTGAAGAAATCAAG 1800
DB 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGATCTCGGCTTGTCTGTGAAGAAATCAAG 1800
QY 1801 CGGAAAGCGACAGAAATGTCAGCATAAACTGTTGAAGCTTAAGGATCAAGGAAAGCTTGAA 1860
DB 1801 CGGAAAGCGACAGAAATGTCAGCATAAACTGTTGAAGCTTAAGGATCAAGGAAAGCTTGAA 1860
QY 1861 GTGGAGAAATATGCGAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAAATTCAGGAG 1920
DB 1861 GTGGAGAAATATGCGAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAAATTCAGGAG 1920
QY 1921 CTCCAAGAGAACTTGGAGAGGCTT----- 1944
DB 1921 CTCCAAGAGAACTTGGAGAGGCTT----- 1944
QY 1945 -----GCAAAAGGAGCGAGCCGAGAGGAGCTGGAAGAGCTGCAGAAACCGAGAG 1992
DB 1945 -----GCAAAAGGAGCGAGCCGAGAGGAGCTGGAAGAGCTGCAGAAACCGAGAG 1992
QY 1993 AATATCCCGCAGGCAAGGAGCGAGCCGAGAGGAGCTGGAAGAGCTGCAGAAACCGAGAG 2040
DB 1993 AATATCCCGCAGGCAAGGAGCGAGCCGAGAGGAGCTGGAAGAGCTGCAGAAACCGAGAG 2040
QY 2052 GATCTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCGCTTCTCTG 2052
DB 2052 GATCTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCGCTTCTCTG 2052
QY 2100 GATCTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCGCTTCTCTG 2100
DB 2100 GATCTCTTGAAGGATCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCGCTTCTCTG 2100
QY 2112 GAGAAACAGGTTAAAGAGACTAGAGACCATGAGCCTAGAGAAACAGACTGAAGGATGAC 2112
DB 2112 GAGAAACAGGTTAAAGAGACTAGAGACCATGAGCCTAGAGAAACAGACTGAAGGATGAC 2112
QY 2160 GAGAAACAGGTTAAAGAGCTTAGAGACCATGAGCCTAGAGAAACAGACTGAAGGATGAC 2160
DB 2160 GAGAAACAGGTTAAAGAGCTTAGAGACCATGAGCCTAGAGAAACAGACTGAAGGATGAC 2160
QY 2172 ATCCAGACAAAATCCCAACAGATCCAGCAGATGCTGATAAAATTTCTGGAGCTCGAAGAG 2172
DB 2172 ATCCAGACAAAATCCCAACAGATCCAGCAGATGCTGATAAAATTTCTGGAGCTCGAAGAG 2172
QY 2232 AAAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTGAGAGTGCACCTGAAACAGAAAGAG 2232
DB 2232 AAAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTGAGAGTGCACCTGAAACAGAAAGAG 2232
QY 2280 AAAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTGAGAGTGCACCTGAAACAGAAAGAG 2280
DB 2280 AAAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTGAGAGTGCACCTGAAACAGAAAGAG 2280
QY 2292 CAGCACTATGAGGAAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2292
DB 2292 CAGCACTATGAGGAAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2292
QY 2340 CAGCACTATGAGGAAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340
DB 2340 CAGCACTATGAGGAAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340
QY 2352 AAGGAGACACTGGAAGACATGATGAGAGACACAGAGAGGAGGCCCATGAGAGGCGAAA 2352
DB 2352 AAGGAGACACTGGAAGACATGATGAGAGACACAGAGAGGAGGCCCATGAGAGGCGAAA 2352
QY 2400 AAGGAGACACTGGAAGACATGATGAGAGACACAGAGAGGAGGCCCATGAGAGGCGAAA 2400
DB 2400 AAGGAGACACTGGAAGACATGATGAGAGACACAGAGAGGAGGCCCATGAGAGGCGAAA 2400
QY 2412 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGGAA 2412
DB 2412 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGGAA 2412
QY 2472 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATCTGAGCAAAATAGCAGCTTTTTTACC 2472
DB 2472 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATCTGAGCAAAATAGCAGCTTTTTTACC 2472
QY 2520 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATCTGAGCAAAATAGCAGCTTTTTTACC 2520
DB 2520 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATCTGAGCAAAATAGCAGCTTTTTTACC 2520
QY 2532 CAAAGGAACTGAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
DB 2532 CAAAGGAACTGAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
QY 2580 CAAAGGAACTGAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
DB 2580 CAAAGGAACTGAGGCGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
QY 2592 CTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2592
DB 2592 CTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2592
QY 2640 AAGATCAGCCCAAGACACACAGTGAAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2640
DB 2640 AAGATCAGCCCAAGACACACAGTGAAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2640
QY 2652 AAGATCAGCCCAAGACACACAGTGAAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2652
DB 2652 AAGATCAGCCCAAGACACACAGTGAAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2652
QY 2700 CAGGAGCTCAGTCTAGAGCAGGAGGAGCAGAACTGAGCTCAAGCGCCAGCTCAGAG 2700
DB 2700 CAGGAGCTCAGTCTAGAGCAGGAGGAGCAGAACTGAGCTCAAGCGCCAGCTCAGAG 2700
QY 2712 CAGGAGCTCAGTCTAGAGCAGGAGGAGCAGAACTGAGCTCAAGCGCCAGCTCAGAG 2712
DB 2712 CAGGAGCTCAGTCTAGAGCAGGAGGAGCAGAACTGAGCTCAAGCGCCAGCTCAGAG 2712
QY 2772 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACACCCCTGACAGCTGCACGGGG 2772
DB 2772 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACACCCCTGACAGCTGCACGGGG 2772
QY 2820 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACACCCCTGACAGCTGCACGGGG 2820
DB 2820 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACACCCCTGACAGCTGCACGGGG 2820
QY 2832 GCGCTGAGAGCCAGCTTCCCGCAGGCGGAGACAGAGCTGGAAGAGACCAAGAGAGCT 2832
DB 2832 GCGCTGAGAGCCAGCTTCCCGCAGGCGGAGACAGAGCTGGAAGAGACCAAGAGAGCT 2832

Db ||||| 2821 GCGCTGGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2880
Qy ||||| 2833 GAAGGAGATCCAGCACTCAGGCACATACAGATGAATCCAGCGCAATTTTCATGCT 2892
Db ||||| 2881 GAAGGAGATCCAGGCACTCAGGCACATACAGATGAATCCAGCGCAATTTTCATGCT 2940
Qy ||||| 2893 CTTTCGTAACAGCTGTACTGTAAATCAGACCTCGGAGGAGCAGCTAAACAGCTGACCGAG 2952
Db ||||| 2941 CTTTCGTAACAGCTGTACTGTAAATCAGACCTCGGAGGAGCAGCTAAACAGCTGACCGAG 3000
Qy ||||| 2953 GACACGCTGAATCMAACACCAAACTTCTACTTGTTCMAACCACTCGATGAGCTTCT 3012
Db ||||| 3001 GACACGCTGAATCMAACACCAAACTTCTACTTGTTCMAACCACTCGATGAGCTTCT 3060
Qy ||||| 3013 GCGCCCAACGAGAGATTCTACAACTCGCAAGTGAAGTGGACCACTCGCGCGGAGATC 3072
Db ||||| 3061 GCGCCCAACGAGAGATTCTACAACTCGCAAGTGAAGTGGACCACTCGCGCGGAGATC 3120
Qy ||||| 3073 AGGAACGAGAGATGAGCTTACAGCCGAGAAGCAAAAGATGGAGGCTCTGAAGACCAAG 3132
Db ||||| 3121 AGGAACGAGAGATGAGCTTACAGCCGAGAAGCAAAAGATGGAGGCTCTGAAGACCAAG 3180
Qy ||||| 3133 TGCACCATCTCGAGGAAACAGTCTATGGATTGGAGGCCCTAAACGATGAGCTGTAGAA 3192
Db ||||| 3181 TGCACCATCTCGAGGAAACAGTCTATGGATTGGAGGCCCTAAACGATGAGCTGTAGAA 3240
Qy ||||| 3193 AAAGAGCGCAGTGGAGGCGCTGGAGGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3252
Db ||||| 3241 AAAGAGCGCAGTGGAGGCGCTGGAGGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
Qy ||||| 3253 TGTTCGGTTTCGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGGCC 3312
Db ||||| 3301 TGTTCGGTTTCGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGGCC 3360
Qy ||||| 3313 GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGGCAGTGAAGAGCACAAGGCT 3372
Db ||||| 3361 GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGGCAGTGAAGAGCACAAGGCT 3420
Qy ||||| 3373 GAGATTCTCGCTCGCAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3432
Db ||||| 3421 GAGATTCTCGCTCGCAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3480
Qy ||||| 3433 GACAACTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3492
Db ||||| 3481 GACAACTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3540
Qy ||||| 3493 CAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGCTTCTGGAAAGCAAGCCAAA 3552
Db ||||| 3541 CAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGCTTCTGGAAAGCAAGCCAAA 3600
Qy ||||| 3553 TTACAGCAGAGTGAACCTGAGAGAAATCAATTTCCGCTGACTCAAGGACTGCA 3612
Db ||||| 3601 TTACAGCAGAGTGAACCTGAGAGAAATCAATTTCCGCTGACTCAAGGACTGCA 3660
Qy ||||| 3613 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGAGGTATCAGCTG 3672
Db ||||| 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGAGGTATCAGCTG 3720
Qy ||||| 3673 GAAAAATTCAGTTCTTCTATTCTCATGAAGGTGAATGAGAGGCACTATTCTCAA 3732
Db ||||| 3721 GAAAAATTCAGTTCTTCTATTCTCATGAAGGTGAATGAGAGGCACTATTCTCAA 3780
Qy ||||| 3733 CAAACCAAACTCATTTGATTTCTGCAAGCCAAAATGGACCAACCTGCTAAAAAGAAAAG 3792
Db ||||| 3781 CAAACCAAACTCATTTGATTTCTGCAAGCCAAAATGSAACCACTGCTAAAAAGAAA-- 3838
Qy ||||| 3793 GGTTTATTTAGTCGACGGAAGAGACCCTGCTTTACCACACAGAGTTCCCTCTGAGTAC 3852
Db ||||| 3839 -----AGGTTCTCTGAGTAC 3855
Qy ||||| 3853 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCGAGCTAGAGGAGCC 3912

Db ||||| 3856 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCGAGCTAGAGGAAGCC 3915
Qy ||||| 3913 CTTTCAGAAAGACCCGCACTCGAGCTCCGCTCCGCGGAGGAAGCTGCCACCGCAAGACA 3972
Db ||||| 3916 CTTTCAGAAAGACCCGCACTCGAGCTCCGCTCCGCGGAGGAAGCTGCCACCGCAAGACA 3975
Qy ||||| 3973 ACGACCAACCAACACCCATCCAGCCAGCAACCGCGAGGAGCAGATCGCCATGTCGCC 4032
Db ||||| 3976 ACGACCAACCAACACCCATCCAGCCAGCAACCGCGAGGAGCAGATCGCCATGTCGCC 4035
Qy ||||| 4033 ATCGTGGGTCGCCAGAGCAACCGCCAGTGCATAGCCTGTGGGCCCGCCCATCCAGC 4092
Db ||||| 4036 ATCGTGGGTCGCCAGAGCAACCGCCAGTGCATAGCCTGTGGGCCCGCCCATCCAGC 4095
Qy ||||| 4093 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGCTTTAAGGAAACGCAATGC 4152
Db ||||| 4096 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGCTTTAAGGAAACGCAATGC 4155
Qy ||||| 4153 CACAAATATCTCACCGATTCAAGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4212
Db ||||| 4156 CACAAATATCTCACCGATTCAAGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4215
Qy ||||| 4213 TGTCTGGATACCGTGCACTTTGGACGCGCAGGCAATCCAAATGTCTCGAATGTCAGGTGATG 4272
Db ||||| 4216 TGTCTGGATACCGTGCACTTTGGACGCGCAGGCAATCCAAATGTCTCGAATGTCAGGTGATG 4275
Qy ||||| 4273 TGTCAACCCCAAGTGTCAAGTGTGCGAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4332
Db ||||| 4276 TGTCAACCCCAAGTGTCAAGTGTGCGAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4335
Qy ||||| 4333 ACACATTTACCGAGGCTTCTGCGGTGACAAATGAACCTCCCGAGGTCTCCAGACCAAG 4392
Db ||||| 4336 ACACATTTACCGAGGCTTCTGCGGTGACAAATGAACCTCCCGAGGTCTCCAGACCAAG 4395
Qy ||||| 4393 GAGCCAGCAGAGCTTGCACCTGGAAAGGTGATGAAGTGGCCAGAGATAACAAACGA 4452
Db ||||| 4396 GAGCCAGCAGAGCTTGCACCTGGAAAGGTGATGAAGTGGCCAGAGATAACAAACGA 4455
Qy ||||| 4453 GGACAGCAAGGCTGGGACAGAGATGATTTGCTCGAGGGATCAAAAGTCCCTCATTTAT 4512
Db ||||| 4456 GGACAGCAAGGCTGGGACAGAGATGATTTGCTCGAGGGATCAAAAGTCCCTCATTTAT 4515
Qy ||||| 4513 GACAAATGAAGCAGAGAGCTGGACAGAGCCGCTGGAAAGATTTGAGCTGTGCTTCCC 4572
Db ||||| 4516 GACAAATGAAGCAGAGAGCTGGACAGAGCCGCTGGAAAGATTTGAGCTGTGCTTCCC 4575
Qy ||||| 4573 GACGGGATGTATCTATTCATGCTGCGTGGTGTGCTTCCGAACTCGCAATACAGCCAAA 4632
Db ||||| 4576 GACGGGATGTATCTATTCATGCTGCGTGGTGTGCTTCCGAACTCGCAATACAGCCAAA 4635
Qy ||||| 4633 GCA----- 4635
Db ||||| 4636 GCAGATGTCCTCATACATACTGAAGATGGAATCTCACCCGACACCACTGCTGCGCCGGG 4695
Qy ||||| 4636 ----- 4636
Db ||||| 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGGCTGGTCAAGCCCTTA 4755
Qy ||||| 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662
Db ||||| 4756 GAATCATGTTGCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAACTG 4815
Qy ||||| 4663 CTTTGGAAACTCCCTGCTGAAACTGGAAAGTGAACCGTCTAGACATGAAGTCAAGCCTG 4722
Db ||||| 4816 CTTTGGAAACTCCCTGCTGAAACTGGAAAGTGAACCGTCTAGACATGAAGTCAAGCCTG 4875
Qy ||||| 4723 CCCTTCAGTGACAGGTTGTTGGTGGGACCGAGGAAGGCTCTACGCGCTGAATGTC 4782
Db ||||| 4876 CCCTTCAGTGACAGGTTGTTGGTGGGACCGAGGAAGGCTCTACGCGCTGAATGTC 4935
Qy ||||| 4783 TTTGAAAAACTCCCTTAACCCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTTATTTATC 4842
Db ||||| 4936 TTTGAAAAACTCCCTTAACCCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTTATTTATC 4995

QY	4843	ARGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTCTGTGGAC	4902	QY	5923	CAAGACAGCAGCAGGCGCGCTCTCGGGAGCCCTGAGGACCCCGCTGTCCAGGTG	5982
DB	4996	ARGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTCTGTGGAC	5055	DB	6076	GAAGACAGCAGCAGGCGCGCTCTCGGGAGCCCGTGTCCAGGTG	6135
QY	4903	GTGAAGAAAGTGAACAGTCTCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA	4962	QY	5983	AACAAGGAGAGGCGAGGTGC	6005
DB	5056	GTGAAGAAAGTGAACAGTCTCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA	5115	DB	6136	AACAAGGTCTGGGACCATCTTC	6158
QY	4963	CCCAACATTTTGAAGCTGTCAAGGCTGCCATCTGTTGGGCGAGGAGATTGAGAC	5022	RESULT 6			
DB	5116	CCCAACATTTTGAAGCTGTCAAGGCTGCCATCTGTTGGGCGAGGAGATTGAGAC	5175	ID	ABQ78870	standard; cDNA; 6165 BP.	
QY	5023	GGGCTCTGCATCTGTGAGCAGTCCGAGCAAGTCTCATTTCTCGCTACACGAAAC	5082	XX	ABQ78870;		
DB	5176	GGGCTCTGCATCTGTGAGCAGTCCGAGCAAGTCTCATTTCTCGCTACACGAAAC	5235	XX	10-OCT-2002 (first entry)		
QY	5083	CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGGCCCTGAGCTGTATCCAC	5142	XX	Human kinase cDNA #1.		
DB	5236	CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGGCCCTGAGCTGTATCCAC	5295	XX	Human; kinase; enzyme; serine-threonine kinase; neotropic; cytostatic;		
QY	5143	TTCAACCAATATACAGTATCTCTTGAACCAATAAATTTCTAGAAATCGACATGAAGCAG	5202	KW	Citron rho-interacting kinase; gene therapy; mental disorder; cancer;		
DB	5296	TTCAACCAATATACAGTATCTCTTGAACCAATAAATTTCTAGAAATCGACATGAAGCAG	5355	KW	gene; ss.		
QY	5203	TACACGCTCAGGAATTCCTGGATAAGATGACATTCCTTGCCACCTGCTGTGTTGCC	5262	XX	Homo sapiens.		
DB	5356	TACACGCTCAGGAATTCCTGGATAAGATGACATTCCTTGCCACCTGCTGTGTTGCC	5415	FH	Key	Location/Qualifiers	
QY	5263	GCCTCTTCAACAGCTTCCTGTCTCAATCTGTCAGTGAACAGCGCAGGCGAGGAG	5322	FT	CDS	1. 6165	
DB	5416	GCCTCTTCAACAGCTTCCTGTCTCAATCTGTCAGTGAACAGCGCAGGCGAGGAG	5475	FT	variation	/tag= a	
QY	5323	GAGTACTTGTCTGTGTTTCAACGAATTTGGAGTCTTCTGGAATCTTACGGAAGAGCTAGC	5382	FT	variation	/product= "Kinase"	
DB	5476	GAGTACTTGTCTGTGTTTCAACGAATTTGGAGTCTTCTGGAATCTTACGGAAGAGCTAGC	5535	FT	variation	/replace(5218,G)	
QY	5383	CGACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTGGCTTACAGAGAACCTAT	5442	FT	variation	/tag= b	
DB	5536	CGACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTGGCTTACAGAGAACCTAT	5595	FT	variation	/standard_name= "Single nucleotide polymorphism"	
QY	5443	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGCAGCTCTCAGCA	5502	FT	variation	/replace(6065,G)	
DB	5596	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGCAGCTCTCAGCA	5655	FT	variation	/tag= b	
QY	5503	GGGACCCCTGCCGAGGCTACCTTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAATT	5562	PN	WO200259325-A2.	/standard_name= "Single nucleotide polymorphism"	
DB	5656	GGGACCCCTGCCGAGGCTACCTTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAATT	5715	XX	01-AUG-2002.		
QY	5563	TCCTCAGGAGCGATTACTTGGGTCTCTCATACGAGTAAATTAAGGCTCATTTGCTGC	5622	XX	20-DEC-2001; 2001WO-US050497.		
DB	5716	TCCTCAGGAGCGATTACTTGGGTCTCTCATACGAGTAAATTAAGGCTCATTTGCTGC	5775	XX	27-DEC-2000; 2000US-0258335P.		
QY	5623	AAGGGAACCTCGTGAAGAGTCCGAGCTGAACACACCGGGGCCCGTCCACCTCCGC	5682	XX	(LEXI-) LEXICON GENETICS INC.		
DB	5776	AAGGGAACCTCGTGAAGAGTCCGAGCTGAACACACCGGGGCCCGTCCACCTCCGC	5835	XX	Yu X, Miranda M, Fiddle CJ;		
QY	5683	AGCAGCCCAACAGCGGCGCCACCCAGTCAACAGCAGCATCAACGAGCGGTGGCC	5742	XX	WPI; 2002-599796/64.		
DB	5836	AGCAGCCCAACAGCGGCGCCACCCAGTCAACAGCAGCATCAACGAGCGGTGGCC	5895	XX	P-PSDB; ABB81927.		
QY	5743	TCCAGCCCAACAGCGGCGCCACCCAGTCAACGAGCAGCATCAACGAGCGGTGGCC	5802	XX	Novel polynucleotide encoding human proteins that are structurally		
DB	5896	TCCAGCCCAACAGCGGCGCCACCCAGTCAACGAGCAGCATCAACGAGCGGTGGCC	5955	XX	similar to animal kinases, useful for drug screening, diagnosis, in gene		
QY	5803	TACCGGAGGGGGAGCGAGTCTCGCAGGAGCAAGTCTCTGCGCGCCCTCGGAGCA	5862	XX	therapy of disorders and diseases e.g. cancer and pharmacogenomic		
DB	5956	TACCGGAGGGGGAGCGAGTCTCGCAGGAGCAAGTCTCTGCGCGCCCTCGGAGCA	6015	XX	applications.		
QY	5863	GAGAAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGGTCCCCCGGAGGCTGTTT	5922	PS	Claim 1; Page 37-39; 50pp; English.		
DB	6016	GAGAAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGGTCCCCCGGAGGCTGTTT	6075	XX	The invention relates to a novel human protein that shares structural		
				CC	similarity with animal kinases, including serine-threonine kinases,		
				CC	particularly Citron rho-interacting kinases. The proteins of the		
				CC	invention have neotropic and cytostatic activity. The polynucleotides may		
				CC	have a use in gene therapy. The encoded novel polypeptides are useful for		
				CC	generating antibodies, as reagents in diagnostic assays, for identifying		
				CC	other cellular gene products related to NHP and as reagents in assays for		
				CC	screening for compounds that are useful in the treatment of mental,		
				CC	biological or medical disorders and diseases including cancer. The		
				CC	sequence encodes a novel human kinase of the invention		
				XX	Sequence 6165 BP; 1735 A; 1550 C; 1679 G; 1201 T; 0 U; 0 Other;		
				SQ	Query Match 91.9%; Score 5661.4; DB 6; Length 6165;		
					Best Local Similarity 95.8%; Pred. No. 0;		

Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

```
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATTCCTTTTGGATGCTGGTCTGCTGAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATTCCTTTTGGATGCTGGTCTGCTGAACCCATT 60

QY 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAACCAACCCCTTTATGACTCAA 120
Db 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAACCAACCCCTTTATGACTCAA 120

QY 121 CAGCAGATGTCCTCTTTTCCGAGAGGGATATTAGATGCCCTCTTTTGTCTTTTGAA 180
Db 121 CAGCAGATGTCCTCTTTTCCGAGAGGGATATTAGATGCCCTCTTTTGTCTTTTGAA 180

QY 181 GAATGAGTCAGCTGCTCTGATGAAGATTAAGCACGTGAGCAACTTTGTCCGGAAGTAT 240
Db 181 GAATGAGTCAGCTGCTCTGATGAAGATTAAGCACGTGAGCAACTTTGTCCGGAAGTAT 240

QY 241 TCCGACACCATAGCTGAGTTTACAGGAGCTCCAGCCTTCGCAAGGACTTCGAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTTACAGGAGCTCCAGCCTTCGCAAGGACTTCGAAGTCAGA 300

QY 301 AGTCTTTGAGTTTGGTCACTTTTGTGTAAGTCAGGTGTTAAGAGAAAAGCAACCGGG 360
Db 301 AGTCTTTGAGTTTGGTCACTTTTGTGTAAGTCAGGTGTTAAGAGAAAAGCAACCGGG 360

QY 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTTATTTGSCCCAGGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTTATTTGSCCCAGGAGCAGGTTTCA 420

QY 421 TTTTGTGGAAGAGCGGCAATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGGAAGAGCGGCAATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480

QY 481 CAGTATGCTTTTTCAGACAAAATCACTTTATCTGATGAGGAAATATCAGCCTGGAGGG 540
Db 481 CAGTATGCTTTTTCAGACAAAATCACTTTATCTGATGAGGAAATATCAGCCTGGAGGG 540

QY 541 GACTTCTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATAAAGCTGATACAGTTT 600
Db 541 GACTTCTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATAAAGCTGATACAGTTT 600

QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCACTGATGGGATACGTGATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCACTGATGGGATACGTGATCGA 660

QY 661 GACATCAAGCCTGAGAAATCTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
Db 661 GACATCAAGCCTGAGAAATCTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720

QY 721 GGATCTGCCGGAATAATGAATCAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
Db 721 GGATCTGCCGGAATAATGAATCAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780

QY 781 CCAGATTATAGCTCTCAAGTGTGATGTTGATGAAGCGGGATGGAAGGCACTTAC 840
Db 781 CCAGATTATAGCTCTCAAGTGTGATGTTGATGAAGCGGGATGGAAGGCACTTAC 840

QY 841 GGCCTGGACTGACTGTGGTTCAGTGGCGGTGATTCGCTATCAGATGATTTTATGCGAGA 900
Db 841 GGCCTGGACTGACTGTGGTTCAGTGGCGGTGATTCGCTATCAGATGATTTTATGCGAGA 900

QY 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACTTTCAATTAACATTAATGAATTCAGCGG 960
Db 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACTTTCAATTAACATTAATGAATTCAGCGG 960

QY 961 TTTTGTGAAATTTCCAGATCACCCCAAGTGAGCAGTACTTTCTTGTATCTGATTCAGAGC 1020
Db 961 TTTTGTGAAATTTCCAGATCACCCCAAGTGAGCAGTACTTTCTTGTATCTGATTCAGAGC 1020

QY 1021 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTTGTGTCATCTTTCTTC 1080
Db 1021 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTTGTGTCATCTTTCTTC 1080
```

```
QY 1081 TCTAAATTTGAGTGGAAACAACATTCGTAACTCTCTCTCCCTCTTCTGTTCCCAACCTCAAG 1140
Db 1081 TCTAAATTTGAGTGGAAACAACATTCGTAACTCTCTCTCCCTCTTCTGTTCCCAACCTCAAG 1140

QY 1141 TCTGACGATGACACCTCCAAATTTTGAATGAACAGAGAAGAAATTCGTGGGTTTCATCTCT 1200
Db 1141 TCTGACGATGACACCTCCAAATTTTGAATGAACAGAGAAGAAATTCGTGGGTTTCATCTCT 1200

QY 1201 CCGTCCAGCTGAGCCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGTTTTTCG 1260
Db 1201 CCGTCCAGCTGAGCCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGTTTTTCG 1260

QY 1261 TACAGCAAGGCACTGGGGATTCCTTTGGTAGATCTGAGTCTGTGTGTCGGGCTGAGCTCC 1320
Db 1261 TACAGCAAGGCACTGGGGATTCCTTTGGTAGATCTGAGTCTGTGTGTCGGGCTGAGCTCC 1320

QY 1321 CCTGCAAGACTAGCTCCATGGAAGAAAGAACTTCTCTCAAAACAAAGAGCTCAAGAC 1380
Db 1321 CCTGCAAGACTAGCTCCATGGAAGAAAGAACTTCTCTCAAAACAAAGAGCTCAAGAC 1380

QY 1381 TCTCAGGCAAGTGTCAAGATCGAGCAAGAAATGACCCGGTTTACATCGGAGAGTGTCA 1440
Db 1381 TCTCAGGCAAGTGTCAAGATCGAGCAAGAAATGACCCGGTTTACATCGGAGAGTGTCA 1440

QY 1441 GAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500
Db 1441 GAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500

QY 1501 TCCCTCTCGAGCAGGAGCTTTGCTTACCTACATCAAGAAATGACAGTAGCTTAAAGCGAAGT 1560
Db 1501 TCCCTCTCGAGCAGGAGCTTTGCTTACCTACATCAAGAAATGACAGTAGCTTAAAGCGAAGT 1560

QY 1561 TTGAGCAAGCACCGATGGAGGTCTCCAGAGAGATGACAAAGCACTTCGACCTTCTCCAT 1620
Db 1561 TTGAGCAAGCACCGATGGAGGTCTCCAGAGAGATGACAAAGCACTTCGACCTTCTCCAT 1620

QY 1621 GATATCAGAGAGCAGCGCGGAAGCTCCAGAAATCAAGAGCAGGAGTACACAGGCTCAA 1680
Db 1621 GATATCAGAGAGCAGCGCGGAAGCTCCAGAAATCAAGAGCAGGAGTACACAGGCTCAA 1680

QY 1681 GTGAAAGAAATGAGGTTGATGAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740
Db 1681 GTGAAAGAAATGAGGTTGATGAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740

QY 1741 CGGAGTGTCTCTACGAATCTGAGTCTGAGAGAGTCTCGGCTTCTGCTGAGGATTCAG 1800
Db 1741 CGGAGTGTCTCTACGAATCTGAGTCTGAGAGAGTCTCGGCTTCTGCTGAGGATTCAG 1800

QY 1801 CGGAAAGCGACAGAAATCTCAGCATAAACTGTGTAAGGCTAAGGATCAAGGGAAGCCTGAA 1860
Db 1801 CGGAAAGCGACAGAAATCTCAGCATAAACTGTGTAAGGCTAAGGATCAAGGGAAGCCTGAA 1860

QY 1861 GTGGGAAATATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1861 GTGGGAAATATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920

QY 1921 CTCGAGAGAAACTTGGAGAAAGGCT----- 1944
Db 1921 CTCGAGAGAAACTTGGAGAAAGGCT----- 1944

QY 1945 -----GCAAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAAGCCAGCAGGAGCCAGCTGCTCGAG 1980
Db 1945 -----GCAAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAAGCCAGCAGGAG 1980

QY 1981 AATATCGCCAGGCAAGAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAAGCCAGAG 2040
Db 1981 AATATCGCCAGGCAAGAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAAGCCAGAG 2040

QY 2052 GATTCTTCTGAAGGCTACAGAAAGAGCTGGTGGAGAGCTGAGGAAAGCGCCCAATCTCTG 2052
Db 2052 GATTCTTCTGAAGGCTACAGAAAGAGCTGGTGGAGAGCTGAGGAAAGCGCCCAATCTCTG 2052

QY 2100 GAGAACAGGTTAAGAGACTTAGAGACCATGAGCGCTAGAGAAACAGACTGAAGAGTAC 2112
Db 2100 GAGAACAGGTTAAGAGACTTAGAGACCATGAGCGCTAGAGAAACAGACTGAAGAGTAC 2112
```

2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2172
2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2220
2173 AAACATCGGAGGCCCAAGTCTCAGCCCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2232
2221 AAACATCGGAGGCCCAAGTCTCAGCCCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2280
2233 CAGCACTATGAGAAAGAGTTAAAGTGTGGCAATCAGATTAAGAAAGACCTGCTGAC 2292
2281 CAGCACTATGAGAAAGAGTTAAAGTGTGGCAATCAGATTAAGAAAGACCTGCTGAC 2340
2293 AAGGAGACACTGAGAAACATGATCGAGACACAGAGGAGGCCCATGAGAAGGGCAAA 2352
2341 AAGGAGACACTGAGAAACATGATCGAGACACAGAGGAGGCCCATGAGAAGGGCAAA 2400
2353 ATTCTAGCGAAACAGAAAGGAGTGAATGCTATGGATTCCAGATCAGATCCCTGGAA 2412
2401 ATTCTAGCGAAACAGAAAGGAGTGAATGCTATGGATTCCAGATCAGATCCCTGGAA 2460
2413 CAGAGGATTGTGAACTGTCTGAAGCCAAATAAATTTGCAGCAATAGCAGTCTTTTACC 2472
2461 CAGAGGATTGTGAACTGTCTGAAGCCAAATAAATTTGCAGCAATAGCAGTCTTTTACC 2520
2473 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2532
2521 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2580
2533 CTGGAGACACAGCTGGAGTTTGGAGGCCCAAGCCGAAACTCGAGGAGCAGCTGGAG 2592
2581 CTGGAGACACAGCTGGAGTTTGGAGGCCCAAGCCGAAACTCGAGGAGCAGCTGGAG 2640
2593 AAGATCAGGCCAACAGACACACAGTGAAGAAATCGGCTGCTGAACTCGAGCAACAGAAATTG 2652
2641 AAGATCAGGCCAACAGACACACAGTGAAGAAATCGGCTGCTGAACTCGAGCAACAGAAATTG 2700
2653 CGGAGGTGAGTCTAGAGACAGAGGAGCAAACTGGAGCTGAAGCCGAGCTCAGAG 2712
2701 CGGAGGTGAGTCTAGAGACAGAGGAGCAAACTGGAGCTGAAGCCGAGCTCAGAG 2760
2713 CTACAGTCTCCCTGCAAGGAGCGAGTCAAGTTCAGACGCTGCAAGGCTGCAAGGCG 2772
2761 CTACAGTCTCCCTGCAAGGAGCGAGTCAAGTTCAGACGCTGCAAGGCTGCAAGGCG 2820
2773 GCCTTGAGAGCAGGCTTCCCAAGGCGAAGACAGAGCTGGAAGACACACAGCAAGCT 2832
2821 GCCTTGAGAGCAGGCTTCCCAAGGCGAAGACAGAGCTGGAAGACACACAGCAAGCT 2880
2833 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAAATCCAGCGCAATTTGATGCT 2892
2881 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAAATCCAGCGCAATTTGATGCT 2940
2893 CTTCTGAACAGCTGTACTGTAAATCAGACACTCGAGAGCAGCTAAACAGCTGACCGAG 2952
2941 CTTCTGAACAGCTGTACTGTAAATCAGACACTCGAGAGCAGCTAAACAGCTGACCGAG 3000
2953 GACACGCTGAACCTCAACACCAAACTTCTACTGTCCAAACAACTCGATGAGGCTTCT 3012
3001 GACACGCTGAACCTCAACACCAAACTTCTACTGTCCAAACAACTCGATGAGGCTTCT 3060
3013 GGGCCCAACGACAGATTTGTAACCTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3072
3061 GGGCCCAACGACAGATTTGTAACCTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3120
3073 ACGGAACGAGAGTGCAGCTTACCGCCAGAACCAACGATGGAGGCTCTGAGACCAAG 3132
3121 ACGGAACGAGAGTGCAGCTTACCGCCAGAACCAACGATGGAGGCTCTGAGACCAAG 3180
3133 TGCAACCATGCTGGAGGAACAGGTCATGGAATTTGGAGGCCCTTAACAGATGAGCTGCTAGAA 3192
3181 TGCACCATGCTGGAGGAACAGGTCATGGAATTTGGAGGCCCTTAACAGATGAGCTGCTAGAA 3240
3193 AAAGAGCGGAGTGGGAGGCTGAGAGGAGCTGCTGGGTGATGAGAAATCCCAAGTTTGAG 3252

3241 AAAGAGCGGAGTGGGAGGCTTGAGGAGGCTCTCGGTGATGAGAAATCCCAAGTTTGAG 3300
3253 TGTGGGTTTCGAGAGCTGAGAGAAATGCTGACACCGAGAAAACAGACAGGGCGAGAGCC 3312
3301 TGTGGGTTTCGAGAGCTGAGAGAAATGCTGACACCGAGAAAACAGACAGGGCGAGAGCC 3360
3313 GATCAGCGGATCACCGAGTCTCGCAGGTGGAGTGGCAGTGAAGGAGCAAAAGGCT 3372
3361 GATCAGCGGATCACCGAGTCTCGCAGGTGGAGTGGCAGTGAAGGAGCAAAAGGCT 3420
3373 GAGATTCCTGCTCTGACAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCCTCTCT 3432
3421 GAGATTCCTGCTCTGACAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCCTCTCT 3480
3433 GACAAGCTCAATGACCTGAGAGAAAGCATGCTATGCTTGAATGAATGCCCAAGCTTA 3492
3481 GACAAGCTCAATGACCTGAGAGAAAGCATGCTATGCTTGAATGAATGCCCAAGCTTA 3540
3493 CAGCAGAAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTTCTGAGAGCAAGCCAAA 3552
3541 CAGCAGAAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTTCTGAGAGCAAGCCAAA 3600
3553 TTACAGCAGCAGATGAGACCTGCAAGAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA 3612
3601 TTACAGCAGCAGATGAGACCTGCAAGAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA 3660
3613 GAAGCTCTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGAATTTGGAGTATCAGCTG 3672
3673 GAAACCAATTCAGGTTCTTATCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3732
3721 GAAACCAATTCAGGTTCTTATCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3780
3733 CAAACCAATTCAGTGAATTTTGTCAAGCCAAAATGGAACCACTGTCTAAAAAGAAAAG 3792
3781 CAAACCAATTCAGTGAATTTTGTCAAGCCAAAATGGAACCACTGTCTAAAAAGAAAAG 3838
3793 GGTATTATTAGTCACGGAAAGGACCTGCTTACCCACACAGAGTTCTCTGCACTAC 3852
3839 -----AGTTCTCTGCACTAC 3855
3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3912
3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3915
3913 CTTCAGAAGACCCGATCGAGCTCCGCTCGCCCGGAGAAAGTGCACCCGCAAGCA 3972
3916 CTTCAGAAGACCCGATCGAGCTCCGCTCGCCCGGAGAAAGTGCACCCGCAAGCA 3975
3973 ACGGACCCACACACCCATCCAGCCAGCCACCGGAGGAGCAGATCGCCATGTCGCC 4032
3976 ACGGACCCACACACCCATCCAGCCAGCCACCGGAGGAGCAGATCGCCATGTCGCC 4035
4033 ATCGTGGCGTCCGAGAGCACCCAGTGCATGAGCCTGTGGCCCGCCATCAGC 4092
4036 ATCGTGGCGTCCGAGAGCACCCAGTGCATGAGCCTGTGGCCCGCCATCAGC 4095
4093 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTCTTAAGGAACGATGCAC 4152
4096 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTCTTAAGGAACGATGCAC 4155
4153 CACAATATTCCTACCGATTCACGTAGGACTGAACATGGAGGCAAAAGTGTGTGTG 4212
4156 CACAATATTCCTACCGATTCACGTAGGACTGAACATGGAGGCAAAAGTGTGTGTG 4215
4213 TGTCTGATACGCTGACCTTTGAGCGCCAGCATCCAAATGTCTCGAATGTCAAGTGTATG 4272
4216 TGTCTGATACGCTGACCTTTGAGCGCCAGCATCCAAATGTCTCGAATGTCAAGTGTATG 4275
4273 TGTCAACCCAAAGTGTCTCACGCTGTGCGAGCCACCTGCGGCTTGCCTGTGAATATGCC 4332

Db	4276	TGTCACCCCAAGTCTCCAGTGCTTGGCCAGCCACCTGGCGGCTTGCTGCTGAATATGCC	4335
Qy	4333	ACACACTTCACCGAGGCTTCTGCCGTGACAAATGAACTCCCCAGGTCTCCAGACCAAG	4392
Db	4336	ACACACTTCACCGAGGCTTCTGCCGTGACAAATGAACTCCCCAGGTCTCCAGACCAAG	4395
Qy	4393	GAGCCACAGAGCTTGCACCTGGAAGGCTGATGAAGTCCCAAGGATACAAACGA	4452
Db	4396	GAGCCACAGAGCTTGCACCTGGAAGGCTGATGAAGTCCCAAGGATACAAACGA	4455
Qy	4453	GGACAGCAAGGCTGGACAGGAGTACATTGCTGGAGGATCAAAAGTCCCTCATTTAT	4512
Db	4456	GGACAGCAAGGCTGGACAGGAGTACATTGCTGGAGGATCAAAAGTCCCTCATTTAT	4515
Qy	4513	GACAAATGAAGCAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC	4572
Db	4516	GACAAATGAAGCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC	4575
Qy	4573	GACGGGATGTATCTATTCAATGTCGCTGGTGTCTTCCGAATCGCAAAATACAGCCAAA	4632
Db	4576	GACGGGATGTATCTATTCAATGTCGCTGGTGTCTTCCGAATCGCAAAATACAGCCAAA	4635
Qy	4633	GCN-----	4635
Db	4636	GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACACCTGCTGGCCCGG	4695
Qy	4636	-----	4635
Db	4696	AGAACCTCTACTTGTGTAGCTCCAGCTTCCCTGACAAACAGCGGTGGGTACCCGCTTA	4755
Qy	4636	-----	4662
Db	4756	GAATCAGTTGTGGCAGGTGGGAGAGTTTCTAGGGAANAAGCAGAACTGATCTAACTG	4815
Qy	4663	CTTGGAATCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAACTGCAAGCTG	4722
Db	4816	CTTGGAATCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAACTGCAAGCTG	4875
Qy	4723	CCCTTCAGTGCACAGTGTGTGTGGGCACCGCAGGAAGGCTCTAGCCCTGAAATGTC	4782
Db	4876	CCCTTCAGTGCACAGTGTGTGTGGGCACCGCAGGAAGGCTCTAGCCCTGAAATGTC	4935
Qy	4783	TTGAAAACCTCCCTAAACCATGTCGCCAGAAATGGAGCAATGTTCCAAATTTATATTATC	4842
Db	4936	TTGAAAACCTCCCTAAACCATGTCGCCAGAAATGGAGCAATGTTCCAAATTTATATTATC	4995
Qy	4843	AAGGACCTGGAGAGCTACTCATGATAGCAGAGAGAGCGGCACCTGTCTTCTGGAC	4902
Db	4996	AAGGACCTGGAGAGCTACTCATGATAGCAGAGAGAGCGGCACCTGTCTTCTGGAC	5055
Qy	4903	GTGAGAAAGTGAACAGTCCCTGGCCCAAGTCCCACTGCCTGCCCCAGCCGACATCTCA	4962
Db	5056	GTGAGAAAGTGAACAGTCCCTGGCCCAAGTCCCACTGCCTGCCCCAGCCGACATCTCA	5115
Qy	4963	CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTGGGCAGGCAAGTTGAGAAC	5022
Db	5116	CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTGGGCAGGCAAGTTGAGAAC	5175
Qy	5023	GGGCTCTGATCTGTGACGCCATGCCAGCAAGTCTGTCATTCTCCGCTTACAAACGAAAC	5082
Db	5176	GGGCTCTGATCTGTGACGCCATGCCAGCAAGTCTGTCATTCTCCGCTTACAAACGAAAC	5235
Qy	5083	CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACTCTGAGCCCTGAGCTGTATCCAC	5142
Db	5236	CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACTCTGAGCCCTGAGCTGTATCCAC	5295
Qy	5143	TTACCAAAATCAGATCTCTTGAACCAATAAATTTAGCAATCGCATGAAGCAG	5202
Db	5296	TTACCAAAATCAGATCTCTTGAACCAATAAATTTAGCAATCGCATGAAGCAG	5355
Qy	5203	TACAGCTCGAGAAATCTCTGATGAAGATGCAATCTCTTGGCACCTGCTGTGTTGCC	5262
Db	5356	TACAGCTCGAGAAATCTCTGATGAAGATGCAATCTCTTGGCACCTGCTGTGTTGCC	5415

RESULT 7

AAS06701	ID	AAS06701	standard; cDNA; 6159 BP.
XX	AC	AAS06701;	
XX	DT	12-SEP-2001	(first entry)

Polynucleotide sequence encoding human protein kinase #1.

Human; protein kinase; PTK; STK; cancer; cardiovascular disease; metabolic disorder; immune related disease; neurological disorder; neurodegenerative disorder; inflammatory disorder; infectious disease; reproductive disorder; gene therapy; ss.

Homo sapiens.

OS	Homo sapiens.
XX	
PN	WO200138503-A2.
XX	

```

PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032085.
XX
XX 24-NOV-1999; 99US-0167482P.
XX
XX (SUGB-) SUGEN INC.
XX
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R,
XX Flanagan P, Clary D;
XX
XX WPI; 2001-343950/36.
XX
XX P-PSDB; AAU03501.
XX
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
XX diagnosing and/or treating e.g. cancer, immune, cardiovascular and
XX neuronal-associated diseases, and microbial infections.
XX
XX Example 1; Fig 1; 433pp; English.
XX
XX RAS06701-AA506757 encode for novel human protein kinases #1-57. The novel
XX protein kinases have been identified as members of the tyrosine or
XX serine/threonine kinase (PTK and STK) families. The polynucleotides
XX encoding protein kinases and the polypeptides may be used in the
XX prevention, diagnosis and treatment of diseases associated with
XX inappropriate kinase expression. For example, they may be used to treat
XX cancers (especially cancers of hematopoietic origin), cardiovascular
XX disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
XX immune related diseases (e.g. rheumatoid arthritis), neurological
XX disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
XX Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
XX disease (e.g. HIV) and reproductive disorders (e.g. infertility).
XX Additionally, polynucleotides encoding protein kinases may be used for
XX gene therapy and as DNA probes in diagnostic assays. The protein kinase
XX polypeptides may be used as antigens in the production of antibodies
XX against the protein kinases and in assays to identify modulators of
XX protein kinase expression and activity
XX
XX Sequence 6159 BP; 1732 A; 1549 C; 1680 G; 1198 T; 0 U; 0 Other;
XX
Query Match          91.7%; Score 5650.4; DB 4; Length 6159;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 5943; Conservative 0; Mismatches 11; Indels 246; Gaps 4;
XX
QY 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTCGATGCTGGTCTGCTGAACCCATT 60
DB 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTCGATGCTGGTCTGCTGAACCCATT 60
XX
QY 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 120
DB 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 120
XX
QY 121 CAGCAGATGTCCTCTTTCCGAGAGGGAATATAGATGCCCTCTTCTCTTTGAA 180
DB 121 CAGCAGATGTCCTCTTTCCGAGAGGGAATATAGATGCCCTCTTCTCTTTGAA 180
XX
QY 181 GAATGAGTCAGCTGCTGTGATGAGATTAGCAGTGAACACTTTCGCGGA--G 237
DB 181 GAATGAGTCAGCTGCTGTGATGAGATTAGCAGTGAACACTTTCGCGGAAGTG 240
XX
QY 238 TATTCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTC 297
DB 241 TATTCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTC 300
XX
QY 298 AGAAGTCTTGTAGTGTGTGCTACTTTGCTGAAGTCAGGTGTAAGAGAGAAAGCAACC 357
DB 301 AGAAGTCTTGTAGTGTGTGCTACTTTGCTGAAGTCAGGTGTAAGAGAGAAAGCAACC 360
XX
QY 358 GGGGACATCTATGCTATGAAGTGATGAAGAGGCTTTATTGGCCCGCAGGAGAGGTT 417
DB 361 GGGGACATCTATGCTATGAAGTGATGAAGAGGCTTTATTGGCCCGCAGGAGGTT 420
XX
QY 418 TCATTTTTTGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAA 477

```


Qy	3670	CTGAAAA	CATTTCAGGT	TCTTATTTCTCATGAAAAGGTCAAAATGAAGGCACATAATTTCT	3729
Db	3721	CTGGAAAA	CATTTCAGGT	TCTTATTTCTCATGAAAAGGTCAAAATGAAGGCACATAATTTCT	3780
Qy	3730	CAAACAACC	AATCTCATTTGATTTTTCTGTAAGCCAAAATGGAACCAACTGCTTAAAAAGAAA	3789	
Db	3781	CAACAAACC	AATCTCATTTGATTTTTCTGTAAGCCAAAATGGAACCAACTGCTTAAAAAGAAA	3840	
Qy	3790	AAGGGTTA	TTTAGTCGACGGAAGAAGGACCCCTGCTTTAACCCACACAGGTTCCTCTGCAG	3849	
Db	3841	A-----	-----AGGTTCTCTGCGAG	3855	
Qy	3850	TACAATGAG	CTGAAGCTGGCCCTTGGAAGAGGAGAAAGCTGCTGTGTCAGAGCTAGAGGAA	3909	
Db	3856	TACAATGAG	CTGAAGCTGGCCCTTGGAAGAGGAGAAAGCTGCTGTGTCAGAGCTAGAGGAA	3915	
Qy	3910	GCCCTTCAG	AAGACC CGC CATCGAGCTCCGCTCCGCCCGGGAGGAAGCTGCCACCCGCAAA	3969	
Db	3916	GCCCTTCAG	AAGACC CGC CATCGAGCTCCGCTCCGCCCGGGAGGAAGCTGCCACCCGCAAA	3975	
Qy	3970	GCAAGCACC	ACCCACACCCCATCCAGCCACGACCCGCGAGGAGCAGATCGGCATGTCTC	4029	
Db	3976	GCAAGCACC	ACCCACACCCCATCCAGCCACGACCCGCGAGGAGCAGATCGGCATGTCTC	4035	
Qy	4030	GCCATCGT	CGGTGCGCCAGAGCACGACCCAGTGCATGAGCCTGCTGGCCCCCGCCATCC	4089	
Db	4036	GCCATCGT	CGGTGCGCCAGAGCACGACCCAGTGCATGAGCCTGCTGGCCCCCGCCATCC	4095	
Qy	4090	AGCGCAGAA	AGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAACCGCATG	4149	
Db	4096	AGCGCAGAA	AGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAACCGCATG	4155	
Qy	4150	CACACAAT	ATTTCTTCAACGATTCAACGTAGGACTGNACTGCGAGGCCACAAAGTGTGCT	4209	
Db	4156	CACACAAT	ATTTCTTCAACGATTCAACGTAGGACTGNACTGCGAGGCCACAAAGTGTGCT	4215	
Qy	4210	GTGTGCT	GTGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCCAGGTG	4269	
Db	4216	GTGTGCT	GTGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCCAGGTG	4275	
Qy	4270	ATGTGTCA	CCCCAAGTGTCTCATGTGTTGGCAGCCACCTTGGCGTGTGCCCTGTGCTGATAT	4329	
Db	4276	ATGTGTCA	CCCCAAGTGTCTCATGTGTTGGCAGCCACCTTGGCGTGTGCCCTGTGCTGATAT	4335	
Qy	4330	GCCACACAT	TTCACCGAGGCCTTCTGCCGTGACAAAATGAACTCCCAGAGTCTCCAGACC	4389	
Db	4336	GCCACACAT	TTCACCGAGGCCTTCTGCCGTGACAAAATGAACTCCCAGAGTCTCCAGACC	4395	
Qy	4390	AAGAGCCCA	GACAGCAGCTTGACCTGGAAAGGTGGATGAAGTGTGCCAGGAATACAAA	4449	
Db	4396	AAGAGCCCA	GACAGCAGCAGCTTGACCTGGAAAGGTGGATGAAGTGTGCCAGGAATACAAA	4455	
Qy	4450	CGAGCAGCA	AGGCTGGGACAGGAAGTACATTTCTCTCGAGGGATCAAAAGTCTCTCAT	4509	
Db	4456	CGAGCAGCA	AGGCTGGGACAGGAAGTACATTTCTCTCGAGGGATCAAAAGTCTCTCAT	4515	
Qy	4510	TATGACAAT	GAAAGCCAGAGAGCTGGAACAGAGCCGCTGGAAAGATTTGAGCTGTGCTT	4569	
Db	4516	TATGACAAT	GAAAGCCAGAGAGCTGGAACAGAGCCGCTGGAAAGATTTGAGCTGTGCTT	4575	
Qy	4570	CCGACCGG	GNATGTATCTATTATGTTGTCGCTTCCGAACCTCGCAAAATACAGCC	4629	
Db	4576	CCGACCGG	GNATGTATCTATTATGTTGTCGCTTCCGAACCTCGCAAAATACAGCC	4635	
Qy	4630	AAAGCAGA-	-----	4637	
Db	4636	AAAGCAGAT	GTCCCATACATACTGAAGATGGAATCTCAACCGCACACCACTGTGTGCC	4695	
Qy	4638	-----	-----	4637	
Db	4696	GGGAGAAC	CCCTTCTACTTGTGTAGTCTCCOAGCTTCCCTGTGACAAACAGCGCTGGTCAACGCC	4755	
Qy	4638	-----	-----AAAGCAGAGCTGATGTCTAAA	4659	

[illegible]

QY 2413 CAGAGATTGTGGAACTGTCTGAAGCCAATAAATCTTGACGAAATAGCAGTCTTTTACC 2472
DB 2458 CAGAGATTGTGGAACTGTCTGAAGCCAATAAATCTTGACGAAATAGCAGTCTTTTACC 2517
QY 2473 CAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAAATTTTAC 2532
DB 2518 CAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAAATTTTAC 2577
QY 2533 CTGGACACACAGSCTGGGAGTTGGAGGCCAGAACCGAAATCTGGAGAGCAGCTGGAG 2592
DB 2578 CTGGAGACACAGGCTGGGAGTTGGAGGCCAGAACCGAAATCTGGAGAGCAGCTGGAG 2637
QY 2593 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGSCTCTCGAACTGGAGACAAAGATTG 2652
DB 2638 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGSCTCTCGAACTGGAGACAAAGATTG 2697
QY 2653 CGGGAGTCACTAGACACAGGAGCAGAACTGGAGCTCAAGGCCAGCTCAGAG 2712
DB 2698 CGGGAGTCACTAGACACAGGAGCAGAACTGGAGCTCAAGGCCAGCTCAGAG 2757
QY 2713 CTACAGCTCTCCCTGACGAGCGCAGTCAACAGTTGACAGCCCTGACAGGCTCAAGGGCG 2772
DB 2758 CTACAGCTCTCCCTGACGAGCGCAGTCAACAGTTGACAGCCCTGACGGCTCAAGGGCG 2817
QY 2773 GGCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAAAGCT 2832
DB 2818 GGCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAAAGCT 2877
QY 2833 GAAAGAGAGATCCAGSCTCAAGGCATCAAGGCACATAGAGTGAATCAGCGCAAAATTTGATGCT 2892
DB 2878 GAAAGAGAGATCCAGGCATCAAGGCATCAAGGCACATAGAGTGAATCAGCGCAAAATTTGATGCT 2937
QY 2893 CTTCTGAACAGCTGTAATCAACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952
DB 2938 CTTCTGAACAGCTGTAATCAACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2997
QY 2953 GACAGCTGTAATCAACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAGCTTCT 3012
DB 2998 GACAGCTGTAATCAACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAGCTTCT 3057
QY 3013 GCGGCCAACGAGAGATTGACAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3072
DB 3058 GCGGCCAACGAGAGATTGACAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3117
QY 3073 ACGGAAACGAGATGAGCTTACAGCAGCAGAAAGCAGAAACGATGGAGGCTGTGAAGACACG 3132
DB 3118 ACGGAAACGAGATGAGCTTACAGCAGCAGAAAGCAGAAACGATGGAGGCTGTGAAGACACG 3177
QY 3133 TGCACCATCTCGAGGAAACAGTCTATGGATTGGAGGCCCTAAACGATGAGCTGTAGAA 3192
DB 3178 TGCACCATCTCGAGGAAACAGTCTATGGATTGGAGGCCCTAAACGATGAGCTGTAGAA 3237
QY 3193 AAAGAGCGCAGTGGAGGCCCTGGAGGAGCGTCTTGGGTGATGAGAAATCCAGTTTGAG 3252
DB 3238 AAAGAGCGCAGTGGAGGCCCTGGAGGAGCGTCTTGGGTGATGAGAAATCCAGTTTGAG 3297
QY 3253 TGTCCGGTTCGAGAGCTGAGAGATGCTGGACCGAGAAACAGAGCAGGCGAGAGCC 3312
DB 3298 TGTCCGGTTCGAGAGCTGAGAGATGCTGGACCGAGAAACAGAGCAGGCGAGAGCC 3357
QY 3313 GATCAGCGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT 3372
DB 3358 GATCAGCGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT 3417
QY 3373 GAGATTCTCGCTTCGAGAGGCTCTCAAGAGCAGAGCTGAAGGCCAGAGCCCTCTCT 3432
DB 3418 GAGATTCTCGCTTCGAGAGGCTCTCAAGAGCAGAGCTGAAGGCCAGAGCCCTCTCT 3477
QY 3433 GACAGCTCAATGACCTCGAGAGAGAGCATGCTTATGCTTGAATGAATGCCGAGAGCTTA 3492
DB 3478 GACAGCTCAATGACCTCGAGAGAGAGCATGCTTATGCTTGAATGAATGCCGAGAGCTTA 3537
QY 3493 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAAAGACGACCAAA 3552

DB 3558 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAAGACGACCAACAA 3597
QY 3553 TTAACAGAGCAGATGGAGCTTCAGAGAAATCAACATTTTCGTCTGACTCAAGGACTGCAA 3612
DB 3598 TTAACAGAGCAGATGGAGCTTCAGAGAAATCAACATTTTCGTCTGACTCAAGGACTGCAA 3657
QY 3613 GAAGCTCTAGATCCGGCTGTACTCTGAAGACAGAAAGAGTGAATTTGGAGTATCAGCTG 3672
DB 3658 GAAGCTCTAGATCCGGCTGTACTCTGAAGACAGAAAGAGTGAATTTGGAGTATCAGCTG 3717
QY 3673 GAAACATTTCAAGTCTCTATTTCTCATGAAAGTGAAATGGAAAGCAGCTATTTCTCAA 3732
DB 3718 GAAACATTTCAAGTCTCTATTTCTCATGAAAGTGAAATGGAAAGCAGCTATTTCTCAA 3777
QY 3733 CAACCCAACTCATTTGATTTTCTGACAGCCAAATGACCAACCTGCTTAAAGACAAAG 3792
DB 3778 CAACCCAACTCATTTGATTTTCTGAAAGCAAAATGGACCAACCTGCTTAAAGACAAAG - 3835
QY 3793 GGTTTATTTAGTCAGAGAAAGAGGACCTGCTTTTACCCACACAGGTTCTCTCAGTAC 3852
DB 3836 -----AGTGGCTCTCAGTAC 3852
QY 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTTAGAGAAAGCC 3912
DB 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTTAGAGAAAGCC 3912
QY 3913 CTTCAAGAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCACCGAAAGCA 3972
DB 3913 CTTCAAGAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCACCGAAAGCA 3972
QY 3973 ACGAACCAACCAACCCATCAACCCAGCCACCGAGGAGCAGATCGCCATCTCGCC 4032
DB 3973 ACGAACCAACCAACCCATCAACCCAGCCACCGAGGAGCAGATCGCCATCTCGCC 4032
QY 4033 ATCTGCGGTGCGCAGAGCAGCCAGCTGCGCATGAGCTGCTGGCCCGCCCATCCAGC 4092
DB 4033 ATCTGCGGTGCGCAGAGCAGCCAGCTGCGCATGAGCTGCTGGCCCGCCCATCCAGC 4092
QY 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCCGGCTTTAAGGAACCGATCAC 4152
DB 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCCGGCTTTAAGGAACCGATCAC 4152
QY 4153 CACAATATTTCTCACCAGATTCAACGTAGGACTGAACTGAGGAGCAGAAAGTGTGCTGTG 4212
DB 4153 CACAATATTTCTCACCAGATTCAACGTAGGACTGAACTGAGGAGCAGAAAGTGTGCTGTG 4212
QY 4213 TGTCTGATACCGTGACTTTGACGCCAGGCATCCAAATGTCTGAAATGTCTAGGTGATG 4272
DB 4213 TGTCTGATACCGTGACTTTGACGCCAGGCATCCAAATGTCTGAAATGTCTAGGTGATG 4272
QY 4273 TGTCAACCCCAAGTGTCTCAGCTGCTGCCAGCCACCTGGCGGCTTGCCTGTGAATATGCC 4332
DB 4273 TGTCAACCCCAAGTGTCTCAGCTGCTGCCAGCCACCTGGCGGCTTGCCTGTGAATATGCC 4332
QY 4333 ACACATTTCCCGAGGCTTCTCCGCTGACAAATGAACCTCCCGAGTCTCCAGACCAAG 4392
DB 4333 ACACATTTCCCGAGGCTTCTCCGCTGACAAATGAACCTCCCGAGTCTCCAGACCAAG 4392
QY 4393 GAGCCCAACGAGCAGCTTGCACCTGGAAGGCTGATGAAGTGGCCAGGAAATAACAAACGA 4452
DB 4393 GAGCCCAACGAGCAGCTTGCACCTGGAAGGCTGATGAAGTGGCCAGGAAATAACAAACGA 4452
QY 4453 GGACAGCAAGGCTGGGACAGAGATGATTTGCTTGGAGGATCAAAAGTCTCTCATTTAT 4512
DB 4453 GGACAGCAAGGCTGGGACAGAGATGATTTGCTTGGAGGATCAAAAGTCTCTCATTTAT 4512
QY 4513 GACAATGAAGCCAGAGAGCTGGACAGAGGCCGGTGGAAATTTGAGCTGTGCTTCCC 4572
DB 4513 GACAATGAAGCCAGAGAGCTGGACAGAGGCCGGTGGAAATTTGAGCTGTGCTTCCC 4572
QY 4573 GACGGGATGATCTATTTCATGCTGCTGCTTCCGAACTCGCAATACAGCCAAA 4632

Db 4573 GACGGGATGATATCTATTCATGTGCGTGTGGTCTTCGAACTCCGAAATACAGCCAAA 4632
 Qy 4633 GCA----- 4635
 Db 4633 GCAGATGTCCTCATACATACATAGAGATGGAATCTCACCCGACACACCTGCTGGCCCGGG 4692
 Qy 4636 ----- 4635
 Db 4693 AGAACCCCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACCGCTTA 4752
 Qy 4636 -----GAAAAAGCAGAAGCTGATGCTAAACTG 4662
 Db 4753 GAATCAGTTGTCGCGAGTGGGAGATTCTAGGGAAGAGCAGAAGCTGATGCTAAACTG 4812
 Qy 4663 CTTGGAAACTCCCTGTGTGAACCTGGAAGTGTGATGACCGTCTAGACATGAACTGCGAGCTG 4722
 Db 4813 CTTGGAAACTCCCTGTGTGAACCTGGAAGTGTGATGACCGTCTAGACATGAACTGCGAGCTG 4872
 Qy 4723 CCCTTCAGTGACACAGGTGTGTTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGAAATGTC 4782
 Db 4873 CCCTTCAGTGACACAGGTGTGTTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGAAATGTC 4932
 Qy 4783 TTGAAAACTCCCTAACCCATGTCCTCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4842
 Db 4933 TTGAAAACTCCCTAACCCATGTCCTCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4992
 Qy 4843 AAGGACCTGGAGAGCTACTCATGATAGCAGAGAGAGGGGACCTGTCTGTGGAC 4902
 Db 4993 AAGGACCTGGAGAGCTACTCATGATAGCAGAGAGAGGGGACCTGTCTGTGGAC 5052
 Qy 4903 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCTGCCAGCCCGACATCTCA 4962
 Db 5053 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCTGCCAGCCCGACATCTCA 5112
 Qy 4963 CCACAACTTTTGAAGTGTCAAGGCTGCCACTTTGTTGGGCGAGGCAAGATTGAGAAC 5022
 Db 5113 CCACAACTTTTGAAGTGTCAAGGCTGCCACTTTGTTGGGCGAGGCAAGATTGAGAAC 5172
 Qy 5023 GGGCTCTGCATCTGTGAGGCTGCCAGCAAGTGTCTATTCCTGCTACACAGAAAC 5082
 Db 5173 GGGCTCTGCATCTGTGAGGCTGCCAGCAAGTGTCTATTCCTGCTACACAGAAAC 5232
 Qy 5083 CTCAGCAATACTGCAATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAAGTGTATCCAC 5142
 Db 5233 CTCAGCAATACTGCAATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAAGTGTATCCAC 5292
 Qy 5143 TTCACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTCTAGAAATCGACATGAAGCAG 5202
 Db 5293 TTCACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTCTAGAAATCGACATGAAGCAG 5352
 Qy 5203 TACACGCTCGAGGAATTCCTGGATAAGAAATGACCATTCCTTGCCACCTGCTGTGTTGCC 5262
 Db 5353 TACACGCTCGAGGAATTCCTGGATAAGAAATGACCATTCCTTGCCACCTGCTGTGTTGCC 5412
 Qy 5263 GCCTCTTCCAAAGCTTCCCTGTCTCAATCTGCAGGTGAACAGCGCAGGCGAGGAG 5322
 Db 5413 GCCTCTTCCAAAGCTTCCCTGTCTCAATCTGCAGGTGAACAGCGCAGGCGAGGAG 5472
 Qy 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTGTGGATTTTACGGAAGACGTAGC 5382
 Db 5473 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTCTGTGGATTTTACGGAAGACGTAGC 5532
 Qy 5383 CGACACAGACGATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTCTAT 5442
 Db 5533 CGACACAGACGATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTTTGGCCCTTACAGAGAACCTCTAT 5592
 Qy 5443 CTGTTTGTGACCCACATTCACCTCACTCGAAGTAATTTAGATCCAGGACGCTCTCAGCA 5502
 Db 5593 CTGTTTGTGACCCACATTCACCTCACTCGAAGTAATTTAGATCCAGGACGCTCTCAGCA 5652
 Qy 5503 GGGACCCCTGCCGAGGATACCTTGGACATCCCGAACCCCGCTTACTTGGGCCCTCGCATTT 5562
 Db 5653 GGGACCCCTGCCGAGGATACCTTGGACATCCCGAACCCCGCTTACTTGGGCCCTCGCATTT 5712

Qy 5563 TCCTCAGAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTCATTGTCTGC 5622
 Db 5713 TCCTCAGAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTCATTGTCTGC 5772
 Qy 5623 AAGGGAACCTTCGTGAAGGAGTCCGGCACTGAACACCAACCGGGCCCGTCCACCTCCCGC 5682
 Db 5773 AAGGGAACCTTCGTGAAGGAGTCCGGCACTGAACACCAACCGGGCCCGTCCACCTCCCGC 5832
 Qy 5683 AGCAGCCCCAACAAAGAGAGGCCCAACCCACGTAACAGAGCAATCACCAAGCGCTGGCC 5742
 Db 5833 AGCAGCCCCAACAAAGAGAGGCCCAACCCACGTAACAGAGCAATCACCAAGCGCTGGCC 5892
 Qy 5743 TCCAGCCAGCGCCCGCGAGGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCG 5802
 Db 5893 TCCAGCCAGCGCCCGCGAGGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCG 5952
 Qy 5803 TACCGCGAGGGCGGAGCCCGAGCTGCGCAGGGAACAAGTCTCTTGGCCGCCCTTGAGCGA 5862
 Db 5953 TACCGCGAGGGCGGAGCCCGAGCTGCGCAGGGAACAAGTCTCTTGGCCGCCCTTGAGCGA 6012
 Qy 5863 GAGAAGTCTCCCGCCCGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGCGCTGTTT 5922
 Db 6013 GAGAAGTCTCCCGCCCGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGCGCTGTTT 6072
 Qy 5923 GAACACAGCAGCAGGGCGCGCTGCTGCGGAGCCGTGAGGACCCCGCTGTCCAGGTTG 5982
 Db 6073 GAACACAGCAGCAGGGCGCGCTGCTGCGGAGCCGTGAGGACCCCGCTGTCCAGGTTG 6132
 Qy 5983 AACAGGGAAGAGGCGAGAGTGCCTCTCAAGTTTTCAGGTTTAACTGTCACT 6037
 Db 6133 AACAGGTTGGGACCACTCTTCAAGTTTAAATCTCAGCCAGAAAAACCAACTCT 6187

RESULT 9
 ADA05641
 ID ADA05641 standard; cDNA; 6189 BP.
 AC ADA05641;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human NOV1a encoding cDNA SEQ ID NO:1.
 XX
 KW human; NOV1a; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytosolic; nontropic; neuroprotective;
 KW antiparkinsonian; antilipidemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..6162
 FT /*tag= a
 FT /product= "NOV1a"
 XX
 WO2003029424-A2.
 XX
 PD 10-APR-2003.
 XX
 XX 02-OCT-2002; 2002WO-US031373.
 PP
 XX 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 28-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373826P.
PR 19-APR-2002; 2002US-0373884P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
PR 17-MAY-2002; 2002US-0381642P.
PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383811P.
PR 25-JUN-2002; 2002US-0391335P.
PR 01-OCT-2002; 2002US-00262511.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malvankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shency SG;
PI Shimkets RA, Rothenberg ME, Leach MD, Agee XL, Berghs C, Dipippo VA;
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX
XX WPI; 2003-381626/36.
XX P-PSDB; ADA05642.
XX
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
XX preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
XX cancer or dyslipidemia, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX
XX Claim 20; Page 98-99; 586pp; English.
XX
XX The present invention describes NOVX proteins, where X can be 1 to 55
XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
XX described above and a carrier; (2) a kit comprising, in one or more
XX containers, the composition described above; (3) an isolated nucleic acid
XX molecule which encodes a NOVX protein of the invention; (4) a vector
XX comprising the nucleic acid molecule described above; (5) a cell
XX comprising the above vector; (6) an antibody that immunospecifically
XX binds to the polypeptide described above; (7) methods for determining the
XX presence or amount of the above polypeptide or nucleic acid molecule in a
XX sample; (8) methods for determining the presence of or predisposition to
XX a disease associated with altered levels of expression of the above
XX polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
XX method of identifying an agent that binds to the polypeptide described
XX above; (10) a method for identifying a potential therapeutic agent for
XX use in treating a pathology that is related to an aberrant expression or
XX aberrant physiological interactions of the polypeptide; (11) a method of
XX screening for a modulator of activity or of latency or predisposition to
XX a pathology associated with the polypeptide; (12) a method for modulating
XX the activity of the polypeptide described above; (13) methods of treating
XX or preventing a pathology associated with the above polypeptide in a
XX mammal; and (14) a method for producing the above polypeptide. NOVX
XX sequences have antidiabetic, anorectic, antibacterial, virucide,
XX immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
XX and antilipaeamic activities, and can be used in gene therapy. The
XX polypeptide is useful in manufacturing a medicament for treating a
XX syndrome associated with a human disease. The polypeptide or the nucleic
XX acid molecule may be used to diagnose, treat or prevent metabolic
XX disorders such as diabetes or obesity, infections, cachexia, cancer,
XX neurodegenerative disorders such as Alzheimer's disease or Parkinson's
XX disease, immune disorders, haematopoietic disorders and various
XX dyslipidaemias. The nucleic acids can also be used as hybridisation
XX probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence encodes a human NOVX protein from
CC the present invention.
XX
SQ Sequence 6189 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;
Query Match 91.4%; Score 5631; DB 7; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;
QY 1 ATGTTGAAGTTCAAAATATGGAGCGCGGAATCCTTTGGATGCTGCTGTCTGCTGAACCCATT 60
DB 1 ATGTTGAAGTTCAAAATATGGAGCGCGGAATCCTTTGGATGCTGCTGTCTGCTGAACCCATT 60
QY 61 GCCAGCCGGGCTCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
DB 61 GCCAGCCGGGCTCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
QY 121 CAGCAGATGCTCCTCTCTTTCCGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
DB 121 CAGCAGATGCTCCTCTCTTTCCGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGGAAGTAT 240
DB 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGGAAGTAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTGGCAAGAGACTTCGAAAGTCAGA 300
DB 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTGGCAAGAGACTTCGAAAGTCAGA 300
QY 301 AGTCTTGTAGTGTGCTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACCGGG 360
DB 301 AGTCTTGTAGTGTGCTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACCGGG 360
QY 361 GACATCTATGCTATGAAGTGAAGAAAGAGAGCTTTATGGCCAGGAGCAGGTTTCA 420
DB 361 GACATCTATGCTATGAAGTGAAGAAAGAGAGCTTTATGGCCAGGAGCAGGTTTCA 420
QY 421 TTTTTCAGGAAGAGCGGAACATATTATCTCGAAGCAAGCCGCTGGATCCCCCAATTA 480
DB 421 TTTTTCAGGAAGAGCGGAACATATTATCTCGAAGCAAGCCGCTGGATCCCCCAATTA 480
QY 481 CAGTATGCTTTTTCAGGACAAATAATCACTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
DB 481 CAGTATGCTTTTTCAGGACAAATAATCACTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
QY 541 GACTTGCTGTCACCTTTTGAATAGATATAGGACCAAGTATGATGAACACCTGTACAGTTT 600
DB 541 GACTTGCTGTCACCTTTTGAATAGATATAGGACCAAGTATGATGAACACCTGTACAGTTT 600
QY 601 TACCTAGCTCAGCTGATTTTGGCTGTTTCAGAGCGTTTCATCTGATGGGATACGTGATCGA 660
DB 601 TACCTAGCTCAGCTGATTTTGGCTGTTTCAGAGCGTTTCATCTGATGGGATACGTGATCGG 660
QY 661 GACATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
DB 661 GACATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
QY 721 GGATCTCCGCGAAATTAATTCACACAGATGGTGAATGCCAACTCCCGATTGGGACC 780
DB 721 GGATCTCCGCGAAATTAATTCACACAGATGGTGAATGCCAACTCCCGATTGGGACC 777
QY 781 CCAGATTACATGGCTCCTGAAGTGTGCTGATGAACCGGGATGAAAGGACCTTAC 840
DB 778 CCAGATTACATGGCTCCTGAAGTGTGCTGATGAACCGGGATGAAAGGACCTTAC 837
QY 841 GGCTTGACTGTGACTGCTGCTGAGTGGGCTGATTCCTTATGAGATGATTTATGGAGA 900
DB 838 GGCTTGACTGTGACTGCTGCTGAGTGGGCTGATTCCTTATGAGATGATTTATGGAGA 897
QY 901 TCCCCCTTCGAGAGGGAACCTCTGCGAGAACCTTCAATAACATTAATGAATTTCCAGCGG 960
DB 898 TCCCCCTTCGAGAGGGAACCTCTGCGAGAACCTTCAATAACATTAATGAATTTCCAGCGG 957

QY 961 TTTTGGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTCTTCTGATCTGATTCAAAGC 1020
DB 958 TTTTGGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTCTTCTGATCTGATTCAAAGC 1017
QY 1021 TTGTTGTGGCGGCAGAAAGAGAGACTGAAAGTTTGAAGGCTTTTGTGCGCATCTTTCTTTC 1080
DB 1018 TTGTTGTGGCGGCAGAAAGAGAGACTGAAAGTTTGAAGGCTTTTGTGCGCATCTTTCTTTC 1077
QY 1081 TCTAAATTTGACTGGACACACATTCCTGCTGCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1140
DB 1078 TCTAAATTTGACTGGACACACATTCCTGCTGCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1137
QY 1141 TCTGACGATGACACCTCCCAATTTTGTATGAACACAGAGAATTTCTGCGGTTTTCATCCTCT 1200
DB 1138 TCTGACGATGACACCTCCCAATTTTGTATGAACACAGAGAATTTCTGCGGTTTTCATCCTCT 1197
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCGGTTTGTGGGGTTTTCG 1260
DB 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCGGTTTGTGGGGTTTTCG 1257
QY 1261 TACAGCAAGCACTGGGATTTCTGATGAACACAGAGAATTTCTGCGGTTTTCATCCTCT 1320
DB 1258 TACAGCAAGCACTGGGATTTCTGATGAACACAGAGAATTTCTGCGGTTTTCATCCTCT 1317
QY 1321 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCGGTTTGTGGGGTTTTCG 1380
DB 1318 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCGGTTTGTGGGGTTTTCG 1377
QY 1381 TCTGACGATGACACCTCCCAATTTTGTATGAACACAGAGAATTTCTGCGGTTTTCATCCTCT 1440
DB 1378 TCTGACGATGACACCTCCCAATTTTGTATGAACACAGAGAATTTCTGCGGTTTTCATCCTCT 1437
QY 1441 GAGGTGAGGCTGTGCTTGTAGTGAAGAGGAGTGGAGCTGAAAGGCTTCTGAGACTCAGAGA 1500
DB 1438 GAGGTGAGGCTGTGCTTGTAGTGAAGAGGAGTGGAGCTGAAAGGCTTCTGAGACTCAGAGA 1497
QY 1501 TCCTCTCTGAGCAGGACCTTGTCTACTATCAGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1560
DB 1498 TCCTCTCTGAGCAGGACCTTGTCTACTATCAGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1557
QY 1561 TTGAGCAAGCAGGATGAGGCTGCTCCAGAGGAGTGAAGAGCACTGAGGCTTCTTCCAT 1620
DB 1558 TTGAGCAAGCAGGATGAGGCTGCTCCAGAGGAGTGAAGAGCACTGAGGCTTCTTCCAT 1617
QY 1621 GATATCAGAGAGCAGGAGCGGAGCTTCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680
DB 1618 GATATCAGAGAGCAGGAGCGGAGCTTCAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677
QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1678 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737
QY 1741 CGGAGTGTCTTACGAATCTGAGTGTGAGAGTCTCGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 1738 CGGAGTGTCTTACGAATCTGAGTGTGAGAGTCTCGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1797
QY 1801 CGGAGTGTCTTACGAATCTGAGTGTGAGAGTCTCGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 1798 CGGAGTGTCTTACGAATCTGAGTGTGAGAGTCTCGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1857
QY 1861 GTGGGAAATATGCGAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920
DB 1858 GTGGGAAATATGCGAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1917
QY 1921 CTCGAGAGAAATCTGGAGAGGCTG----- 1945
DB 1918 CTCGAGAGAAATCTGGAGAGGCTGTTAAAGCCAGCAGGAGCCAGGCTGTGTCAG 1977
QY 1946 -----CAAGAGAGCAGCGCAGAGGAGCTGGAGAGCTGGAGAGCTGGAGAGCTGGAGAGCTGGAGAG 1992
DB 1978 AATATCGCCAGCAGAGAGCGAGCGGAGAGGAGCTGGAGAGCTGGAGAGCTGGAGAGCTGGAGAG 2037
QY 1993 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGAGCTGGAGAGCGCCCATCTCTG 2052

DB 2038 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGAAACGCCGCCATCTCTG 2097
QY 2053 GAGAAACAGGTAAAGAGACTAGAGACCATTGGAGGCTAGAGAAACAGACAGCTGAGAGGATGAC 2112
DB 2098 GAGAAACAGGTAAAGAGACTAGAGACCATTGGAGGCTAGAGAAACAGACAGCTGAGAGGATGAC 2157
QY 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGGAGCTCGAAGAG 2172
DB 2158 ATCCAGACAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGGAGCTCGAAGAG 2217
QY 2173 AAAATCGGGAGGCCCAGTCTCAGCCAGCAGCCTAGAGTGCACCTGAAACAGAAAGAG 2232
DB 2218 AAAATCGGGAGGCCCAGTCTCAGCCAGCAGCCTAGAGTGCACCTGAAACAGAAAGAG 2277
QY 2233 CAGCACTATGAGGAAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC 2292
DB 2278 CAGCACTATGAGGAAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC 2337
QY 2293 AAGGAGACACTGGAGAACATGATCAGAGACA CAGAGAGAGGCCCATGAGAGGGCAAA 2352
DB 2338 AAGGAGACACTGGAGAACATGATCAGAGACA CAGAGAGAGGCCCATGAGAGGGCAAA 2397
QY 2353 ATTTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCCTAGATCAGATCCTCTCGAA 2412
DB 2398 ATTTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCCTAGATCAGATCCTCTCGAA 2457
QY 2413 CAGAGGATTTGGAACTGTCTGAGGCCAATAAATTTGAGCAAAATAGCAGTCTTTTACC 2472
DB 2458 CAGAGGATTTGGAACTGTCTGAGGCCAATAAATTTGAGCAAAATAGCAGTCTTTTACC 2517
QY 2473 CAAAGGACATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532
DB 2518 CAAAGGACATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577
QY 2533 CTGAGAGACACAGGCTGGAAAGTTGGAGGCCAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2592
DB 2578 CTGAGAGACACAGGCTGGAAAGTTGGAGGCCAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2637
QY 2593 AAGATCAGGCCAACAGACCA CAGTGACAAGAACTCGGCTGTCTGAACTGGAGAGACAAGTTG 2652
DB 2638 AAGATCAGGCCAACAGACCA CAGTGACAAGAACTCGGCTGTCTGAACTGGAGAGACAAGTTG 2697
QY 2653 CGGAGGTCAGTCTAGAGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAAGAG 2712
DB 2698 CGGAGGTCAGTCTAGAGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAAGAG 2757
QY 2713 CTACAGCTCTCTCTGAGGAGCGGAGTCA CAGTTGACGCCCTGAGGCTGCA CCGGCG 2772
DB 2758 CTACAGCTCTCTCTGAGGAGCGGAGTCA CAGTTGACGCCCTGAGGCTGCA CCGGCG 2817
QY 2773 GCTCTGAGAGCAGCTTCGCCAGCGAAGACAGAGCTGGAGAGACCAAGAGACCAAGAGAGCT 2832
DB 2818 GCTCTGAGAGCAGCTTCGCCAGCGAAGACAGAGCTGGAGAGACCAAGAGAGCT 2877
QY 2833 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2892
DB 2878 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2937
QY 2893 CTTCTGTAACAGCTGTACTGTAACTCAGAGCTGGAGAGCAGCTTAACCGAGCTGACCGAG 2952
DB 2938 CTTCTGTAACAGCTGTACTGTAACTCAGAGCTGGAGAGCAGCTTAACCGAGCTGACCGAG 2997
QY 2953 GACAACGCTGAACCTCAACCAACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT 3012
DB 2998 GACAACGCTGAACCTCAACCAACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT 3057
QY 3013 GCGCCCAACGACAGATTTGTAACCTCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3072
DB 3058 GCGCCCAACGACAGATTTGTAACCTCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3117
QY 3073 ACAGAAACGAGAGATGACGCTTACAGCGAGCAAGCAACGATGAGGCTCTGAAGACCAAG 3132

QY 5143 TTACCAATTACAGTATCTCATTTGGAAACCAATAAATTTACGAAATCCGACATGAAGCAG 5202
| | | | |
Db 5293 TTACCAATTACAGTATCTCATTTGGAAACCAATAAATTTACGAAATCCGACATGAAGCAG 5352
| | | | |
QY 5203 TACAGGCTCAGGAAATTCCTGGATAAGAAATGACCATTCCTTGGACATCTGCTGTTGGCC 5262
| | | | |
Db 5353 TACAGGCTCAGGAAATTCCTGGATAAGAAATGACCATTCCTTGGACATCTGCTGTTGGCC 5412
| | | | |
QY 5263 GCCTCTTCCAAACAGCTTCCCTGCTCTCAATCGTGCAGGTGAACAGCGAGGCGAGCAGAG 5322
| | | | |
Db 5413 GCCTCTTCCAAACAGCTTCCCTGCTCTCAATCGTGCAGGTGAACAGCGAGGCGAGCAGAG 5472
| | | | |
QY 5323 GAGTACTTGCTGTGTTTCCACGAAATTTGGAGTGTTCGTGGAATCTTACGGAAGACGTAGC 5382
| | | | |
Db 5473 GAGTACTTGCTGTGTTTCCACGAAATTTGGAGTGTTCGTGGAATCTTACGGAAGACGTAGC 5532
| | | | |
QY 5383 CGCAGACAGCATCTCAAGTGGAGTCGCTTACCTTTTGCCCTTACGAGAACCTAT 5442
| | | | |
Db 5533 CGCAGACAGCATCTCAAGTGGAGTCGCTTACCTTTTGCCCTTACGAGAACCTAT 5592
| | | | |
QY 5443 CTGTTGTGACCACTTCAACTCACTCGAAGTAATTTAGATCCAGCACGCTCTCCAGCA 5502
| | | | |
Db 5593 CTGTTGTGACCACTTCAACTCACTCGAAGTAATTTAGATCCAGCACGCTCTCCAGCA 5652
| | | | |
QY 5503 GGGACCCCTGCCCGAGGTACCTGGACATCCGAAACCGCGCTACCTGGGCGCTGCCATT 5562
| | | | |
Db 5653 GGGACCCCTGCCCGAGGTACCTGGACATCCGAAACCGCGCTACCTGGGCGCTGCCATT 5712
| | | | |
QY 5563 TCCTCAGGACGATTTACTTTGGCGTCTCTATACGAGGATAAATTAAGGGTCAATTTGCTGC 5622
| | | | |
Db 5713 TCCTCAGGACGATTTACTTTGGCGTCTCTATACGAGGATAAATTAAGGGTCAATTTGCTGC 5772
| | | | |
QY 5623 AAGGAAACCTCTGTAAGGAGTCCGACATGAACACACCGGGGCCCTGCACCTCCCGC 5682
| | | | |
Db 5773 AAGGAAACCTCTGTAAGGAGTCCGACATGAACACACCGGGGCCCTGCACCTCCCGC 5832
| | | | |
QY 5683 AGCAGCCCCAACAGGAGGCCACCCACGTCATCAACGAGCACATCAACGAGCGGTGGCC 5742
| | | | |
Db 5833 AGCAGCCCCAACAGGAGGCCACCCACGTCATCAACGAGCACATCAACGAGCGGTGGCC 5892
| | | | |
QY 5743 TCCAGCCAGCGCGCCCGAGGCGCCAGCCACCCCGAGAGCCAAAGACACCCACCGC 5802
| | | | |
Db 5893 TCCAGCCAGCGCGCCCGAGGCGCCAGCCACCCCGAGAGCCAAAGACACCCACCGC 5952
| | | | |
QY 5803 TACCGGAGGCGGACCGAGTCCGAGGACAGTCTCTGGCGCCCGCTGGAGCGA 5862
| | | | |
Db 5953 TACCGGAGGCGGACCGAGTCCGAGGACAGTCTCTGGCGCCCGCTGGAGCGA 6012
| | | | |
QY 5863 GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5922
| | | | |
Db 6013 GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6072
| | | | |
QY 5923 GAGACAGCAGGAGGCGCGGCTGCTGGAGCGGAGCGGTGAGGACCCCGCTGCCAGTG 5982
| | | | |
Db 6073 GAGACAGCAGGAGGCGCGGCTGCTGGAGCGGAGCGGTGAGGACCCCGCTGCCAGTG 6132
| | | | |
QY 5983 AACAGGGAGAGGCGGAGGCTGCTCTCAAGTTTTCACCGTTTAAACACTGTCACTT 6037
| | | | |
Db 6133 AACAGGGTGGGACCACTTCCAGTATTAATCTCAGCCAGAAACCAACTCTCT 6187
| | | | |

RESULT 10

ABS63435

ID ABS63435 standard; cDNA; 6201 BP.

AC ABS63435;

XX ABS63435;

XX 15-NOV-2002 (first entry)

XX Human cDNA, homologous to kinases, designated NOV3a.

XX Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;
XX atherosclerosis; diabetes; cell signalling; metabolic pathway;

KW cellular receptor; downstream effector; cancer; gene therapy;
KW hypertension; congenital heart defect; aortic stenosis; obesity;
KW infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
KW neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
KW haematopoietic disease; scleroderma; fertility; immunogen;
KW idiopathic thrombocytopenic purpura; graft versus host disease;
KW Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
KW systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
KW stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;
KW pain; alcoholism; transgenic.

XX Homo sapiens.

XX WC200226826-A2.

XX 04-APR-2002.

XX 27-SEP-2001; 2001WO-US042336.

XX 27-SEP-2000; 2000US-0235631P.

XX 27-SEP-2000; 2000US-0235633P.

XX 27-SEP-2000; 2000US-0235808P.

XX 27-SEP-2000; 2000US-0236064P.

XX 27-SEP-2000; 2000US-0236065P.

XX 27-SEP-2000; 2000US-0236066P.

XX 28-SEP-2000; 2000US-0236135P.

XX 03-OCT-2000; 2000US-0237434P.

XX 05-OCT-2000; 2000US-0238321P.

XX 06-OCT-2000; 2000US-0238396P.

XX 16-MAR-2001; 2001US-0276657P.

XX 31-MAY-2001; 2001US-0294823P.

XX 12-JUL-2001; 2001US-0304868P.

XX 26-SEP-2001; 2001US-00235631.

(CURA-) CURAGEN CORP.

Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;

Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;

Burgess CB, Padigaru M, Kekuda R, Spytek KA, Leach MD, Shinkets RA;

WPI; 2002-499860/53.

P-PSDB; ABG78362.

Novel isolated NOVX polypeptides and polynucleotides homologous to
attratin, plexin, papin-like family of proteins, useful for treating
atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and
stroke.

Claim 8; Page 40-42; 308pp; English.

The invention discloses the isolated human polypeptides and
polynucleotides encoding them, that have been designated NOVX. The
polypeptides, polynucleotides and antibodies are useful in treating or
preventing a NOVX-associated disorder which is cardiomyopathy,
atherosclerosis and diabetes in a human, where the disorder is related to
cell signal processing and metabolic pathway modulation. They can also be
used in determining the presence of, or predisposition to, a disease
associated with altered levels of the polypeptides and polynucleotides of
any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for
identifying an agent that binds to, or that modulates the expression or
activity of the polypeptide, for identifying an agent which is cellular
receptor or downstream effector, for treating or preventing a NOVX-
associated disorder and as a pharmaceutical composition comprising the
polypeptide, polynucleotide or the antibody. The polypeptides and
polynucleotides are useful in diagnostic applications (e.g. as a marker
for cancerous cells or tissue types) where their amounts are assessed, or
for the manufacture of a medicament (e.g. gene therapy) for treating or
preventing disorders or syndromes such as hypertension, congenital heart
defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,
Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,
haemophilia, dyslipidemias, haematopoietic diseases, scleroderma,
fertility, idiopathic thrombocytopenic purpura, graft versus host

CC	diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
CC	disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
CC	allergy, stroke, anxiety, Lesh-Nyhan syndrome, schizophrenia, cerebellar
CC	ataxia, pain and alcoholism. They may also be used as immunogens to
CC	produce antibodies specific for the invention, and as vaccines.
CC	Transgenic cells containing a NOVX expressing construct are useful to
CC	produce non-human transgenic animals for studying the function and/or
CC	activity of the NOVX proteins and for identifying and/or evaluating
CC	modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC	expressing construct are useful to produce non-human transgenic animals
CC	for studying the function and/or activity of the NOVX proteins and for
CC	identifying and/or evaluating modulators of NOVX protein activity. The
CC	sequences presented in ABS63431-ABS63444 are the human NOV1-NOV8 cDNAs
XX	
SQ	Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;
Query Match 91.4%; Score 5629.8; DB 6; Length 6201;	
Best Local Similarity 95.7%; P: 0; Mismatches 22; Indels 246; Gaps 4;	
Matches 5929; Conservative 0;	
QY	1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTGTCTGCTGAACCCATT 60
Db	1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTGTCTGCTGAACCCATT 60
QY	61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 120
Db	61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 120
QY	121 CAGCAGATGTCCTCTTTCCGAGAGGGAATTTAGATGCCCTTTTGTCTCTTTGAA 180
Db	121 CAGCAGATGTCCTCTTTCCGAGAGGGAATTTAGATGCCCTTTTGTCTCTTTGAA 180
QY	181 GAATGCAGTCACGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT 240
Db	181 GAATGCAGTCACGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT 240
QY	241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGAGCTTCGAAGTCAGA 300
Db	241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGAGCTTCGAAGTCAGA 300
QY	301 AGTCTGTAGGTTGGTGCACCTTTGCTGAAGTGCAGGTGGTAAAGAGAAACCAACCGGG 360
Db	301 AGTCTGTAGGTTGGTGCACCTTTGCTGAAGTGCAGGTGGTAAAGAGAAACCAACCGGG 360
QY	361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTCGGCCAGGAGCAGGTTTCA 420
Db	361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATTCGGCCAGGAGCAGGTTTCA 420
QY	421 TTTTGTGAGAGAGCGGACATATTTATCTCGAAGCAAGCCGCTGGATCCCCCAATTA 480
Db	421 TTTTGTGAGAGAGCGGACATATTTATCTCGAAGCAAGCCGCTGGATCCCCCAATTA 480
QY	481 CAGTATGCTCTTCAGGACAAAATCACTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540
Db	481 CAGTATGCTCTTCAGGACAAAATCACTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540
QY	541 GACTTGTCTACCTTTTGAATAGATAGGACCACTTATGATGAAACCTGATACAGTTT 600
Db	541 GACTTGTCTACCTTTTGAATAGATAGGACCACTTATGATGAAACCTGATACAGTTT 600
QY	601 TACCTAGCTGAGTATTTGGCTGTTTACAGGCTTCACTGATGGGATACGTCATCGA 660
Db	601 TACCTAGCTGAGTATTTGGCTGTTTACAGGCTTCACTGATGGGATACGTCATCGA 660
QY	661 GACATCAAGCCTCAGAACATCTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 720
Db	661 GACATCAAGCCTCAGAACATCTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 720
QY	721 GGATCTCCCGGAATGATTCACACAGATGGTGAATCCCAAACTCCCGATTTGGGACC 780
Db	721 GGATCTCCCGGAATGATTCACACAGATGGTGAATCCCAAACTCCCGATTTGGGACC 777
QY	781 CCAGATTACATGGCTCTGAAAGTGTGCTGACTGTGATGAACGGGGATGGAAGACCTTAC 840

Db	778 CCAGATTACATGGCTCTGAAAGTGTGCTGATGAACGGGGATGAAAAGACCTTAC 837
QY	841 GGCTGAGCTGACCTGCTGAGTGGGCTGATGCTATGAGATGATTTATGGAGA 900
Db	838 GGCTGAGCTGCTGAGTGGGCTGATGCTATGAGATGATTTATGGAGA 897
QY	901 TCCCTCTTCGACAGAGGAACTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 960
Db	898 TCCCTCTTCGACAGAGGAACTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 957
QY	961 TTTTGAATTTCCAGATGATCCCCAAAGTGGAGTACTTTCTGATCTGATTTCAAGC 1020
Db	958 TTTTGAATTTCCAGATGATCCCCAAAGTGGAGTACTTTCTGATCTGATTTCAAGC 1017
QY	1021 TTGTTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1080
Db	1018 TTGTTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1077
QY	1081 TCTAAATTTGACTGGAAACAACTTCTGTTGATGAACACAGAGAAATTCGTGGGTTTCATCTCT 1140
Db	1078 TCTAAATTTGACTGGAAACAACTTCTGTTGATGAACACAGAGAAATTCGTGGGTTTCATCTCT 1137
QY	1141 TCTGACGATGACCTTCCAAATTTTGTGATGAACACAGAGAAATTCGTGGGTTTCATCTCT 1200
Db	1138 TCTGACGATGACCTTCCAAATTTTGTGATGAACACAGAGAAATTCGTGGGTTTCATCTCT 1197
QY	1201 CCGTGCAGCTGAGCCCTCAGCTTCTCGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260
Db	1198 CCGTGCAGCTGAGCCCTCAGCTTCTCGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1257
QY	1261 TACAGCAAGGCTGCGGATTTCTGTTGATGATCTGAGTCTGTTGTCGGGCTTGGACTCC 1320
Db	1258 TACAGCAAGGCTGCGGATTTCTGTTGATGATCTGAGTCTGTTGTCGGGCTTGGACTCC 1317
QY	1321 CCTGCCAAGACTAGCTCCATGGAAGAAACTTCTCTCAAAAAAGCAAGAGCTACAAGAC 1380
Db	1318 CCTGCCAAGACTAGCTCCATGGAAGAAACTTCTCTCAAAAAAGCAAGAGCTACAAGAC 1377
QY	1381 TCTCAGGCAAGTGTCAAGAATGGAGCAGGAATGACCCGGTTTACATCGGAGAGTGTCA 1440
Db	1378 TCTCAGGCAAGTGTCAAGAATGGAGCAGGAATGACCCGGTTTACATCGGAGAGTGTCA 1437
QY	1441 GAGGTGAGGCTGTGCTTAGTTCAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db	1438 GAGGTGAGGCTGTGCTTAGTTCAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497
QY	1501 TCCCTCTCTGAGCAGGACCTTGTCTACATCAAGAAATGAGTAAAGCGAAGT 1560
Db	1498 TCCCTCTCTGAGCAGGACCTTGTCTACATCAAGAAATGAGTAAAGCGAAGT 1557
QY	1561 TTGAGCAAGCAGCGATGAGGCTGCCAGGAGGATGACAAAGCACTGACGCTTCTCCAT 1620
Db	1558 TTGAGCAAGCAGCGATGAGGCTGCCAGGAGGATGACAAAGCACTGACGCTTCTCCAT 1617
QY	1621 GATATCAGAGCAGAGCGGAAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680
Db	1618 GATATCAGAGCAGAGCGGAAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677
QY	1691 GTGGAAGAAATGAGGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAGAGAGA 1740
Db	1678 GTGGAAGAAATGAGGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAGAGAGA 1737
QY	1741 CGGAGTGTATCTCTAGAAATCTGAGCAGTCTCGGCTTGTGCTGCTGAAGAAATTCAG 1800
Db	1738 CGGAGTGTATCTCTAGAAATCTGAGCAGTCTCGGCTTGTGCTGCTGAAGAAATTCAG 1797
QY	1801 CGGAAAGCAGCAGAAATGTGACATAAACTGTGTGAAGGCTAAGGATCAAGGGAGCCTGAA 1860
Db	1798 CGGAAAGCAGCAGAAATGTGACATAAACTGTGTGAAGGCTAAGGATCAAGGGAGCCTGAA 1857
QY	1861 GTGGAGAAATGCGAAACTGGAGAGATCAATCTGAGCAGCAGCTCAAAATTCAGAG 1920

ADA05653

ID ADA05653 standard; cDNA; 6201 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Human NOV1g encoding cDNA SEQ ID NO:13.

human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

immunomodulator; cytosolic; nontropic; neuroprotective;

antiparkinsonian; antilipemic; gene therapy; human disease;

metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.

Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

preventing or treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or dyslipidemia, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 20; Page 104-105; 586pp; English.

The present invention describes NOVX proteins, where X can be 1 to 55 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid molecule which encodes a NOVX protein of the invention; (4) a vector comprising the nucleic acid molecule described above; (5) a cell comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity of or latency of predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide. NOVX sequences have antidiabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, nontropic, neuroprotective, antiparkinsonian and antilipemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a syndrome associated with a human disease. The polypeptide or the nucleic acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence encodes a human NOVX protein from the present invention.

Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;

Query Match 91.4%; Score 5629.8; DB 7; Length 6201;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;

QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAACCAATT 60
DB 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAACCAATT 60
QY 61 GCCAGCGCGGCTCCAGGCTGGAATCTGTTCTCCAGGGGAAACCACTTTATGACTCAA 120
DB 61 GCCAGCGCGGCTCCAGGCTGGAATCTGTTCTCCAGGGGAAACCACTTTATGACTCAA 120
QY 121 CAGCAGATGCTCTCTTTTCCAGAGAGGATATTAGATGCCCTTTTGTCTCTTTGAA 180
DB 121 CAGCAGATGCTCTCTTTTCCAGAGAGGATATTAGATGCCCTTTTGTCTCTTTGAA 180
QY 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAGACACGTGAGCAACTTTTTCGGGAAGTAT 240
DB 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAGACACGTGAGCAACTTTTTCGGGAAGTAT 240
QY 241 TCCGACACCATAGCTGAGTACAGGAGCTCCAGCCTTCGCGAAGGACTTCGAGGTCAGA 300
DB 241 TCCGACACCATAGCTGAGTACAGGAGCTCCAGCCTTCGCGAAGGACTTCGAGGTCAGA 300
QY 301 AGTCTTTAGGTTGTGCTCACCCTTTGCTGAAGTGCAGGTGCTGAAGAGAAACCAACCGG 360
DB 301 AGTCTTTAGGTTGTGCTCACCCTTTGCTGAAGTGCAGGTGCTGAAGAGAAACCAACCGG 360
QY 361 GACATCTATGCTATGAAGTATGAGAGAGGCTTTATTCGCCAGGAGGCTTTCA 420
DB 361 GACATCTATGCTATGAAGTATGAGAGAGGCTTTATTCGCCAGGAGGCTTTCA 420

(CURA-) CURAGEN CORP.

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
Ort T, Gorman L, Zerbunen BD, Anderson DW, Zhong M, Catterton E;
Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
Shinkets RA, Rotherberg ME, Leach WD, Agee MD, Berghs C, Dipippo VA;
Eisen AD, Gangolli EA, Rieger DK, Spaderna SK;

WPI; 2003-381626/36.

P-PSDB; ADA05654.

New NOVX polypeptides and nucleic acids, useful for diagnosing,

Db 361 GACATCTATGCTATGAAAGTGATGAAGAAGGCTTTATTGGCCACGAGCAGGTTTCA 420
QY 421 TTTTGTGAGAGAGCGGCAACATATTTCTCGAAGCACAAGCCGTGGATCCCCAAATTA 480
Db 421 TTTTGTGAGAGAGCGGCAACATATTTCTCGAAGCACAAGCCGTGGATCCCCAAATTA 480
QY 481 CAGTATGCCCTTTCAGGACAAAATCACCCTTTATCTGATGAGGAATATCAGCCTCGAGGG 540
Db 481 CAGTATGCCCTTTCAGGACAAAATCACCCTTTATCTGATGAGGAATATCAGCCTCGAGGG 540
QY 541 GACTTGCTGTCTATTTGTAATAGATGAGGACCAAGTCTAGATGAATACTGATACAGTTT 600
Db 541 GACTTGCTGTCTATTTGTAATAGATGAGGACCAAGTCTAGATGAATACTGATACAGTTT 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCATCTGATGGGATACGTCGATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCATCTGATGGGATACGTCGATCGG 660
QY 661 GACATCAAGCCTGAGAACATTCCTGTTGACCGACAGGACACATCAAGCTGGTGGATTTT 720
Db 661 GACATCAAGCCTGAGAACATTCCTGTTGACCGACAGGACACATCAAGCTGGTGGATTTT 720
QY 721 GGNCTGCGCGGNAATGATTCACACAGATGGTGAATGCCAACTCCCGATTGGGACC 780
Db 721 GGNCTGCGCGGNAATGATTCACACAA--GGTGAATGCCAACTCCCGATTGGGACC 777
QY 781 CCAGATTACATGGCTCCTGAAAGTGTGACTGTGATGAACGGGATGGAAAAGGACCTTAC 840
Db 781 CCAGATTACATGGCTCCTGAAAGTGTGACTGTGATGAACGGGATGGAAAAGGACCTTAC 837
QY 841 GGCCTGGACTGTGACTGGTGGTCACTGGGGGTGATGGCTATGAGATGATTTATGGGAGA 900
Db 841 GGCCTGGACTGTGACTGGTGGTCACTGGGGGTGATGGCTATGAGATGATTTATGGGAGA 897
QY 901 TCCCCCTTCCAGAGGAACTCTGCCAGAACCTTCAATAATATGAATTTCCAGCGG 960
Db 901 TCCCCCTTCCAGAGGAACTCTGCCAGAACCTTCAATAATATGAATTTCCAGCGG 957
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATCTGAATCAAAGC 1020
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATCTGAATCAAAGC 1017
QY 1021 TTGCTGTGGGCGCAGAAAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
Db 1021 TTGCTGTGGGCGCAGAAAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1077
QY 1081 TCTAAAATTGACTGGAACCAATTCGTAACTCTCCCTCCCTTCGTTCCACCTCAAG 1140
Db 1081 TCTAAAATTGACTGGAACCAATTCGTAACTCTCCCTCCCTTCGTTCCACCTCAAG 1137
QY 1141 TCTGACGATGACCTCCAAATTTGATGAACACAGAGAAGATTCGTGGGTTTCATCCTCT 1200
Db 1141 TCTGACGATGACCTCCAAATTTGATGAACACAGAGAAGATTCGTGGGTTTCATCCTCT 1197
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTGCGCTTTTGTGGGTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTGCGCTTTTGTGGGTTTTCG 1257
QY 1261 TACAGCAAGGCACTGGGGATCTTGGTATCTGAGTCTGTTGTGCGGGTCTGGACTCC 1320
Db 1261 TACAGCAAGGCACTGGGGATCTTGGTATCTGAGTCTGTTGTGCGGGTCTGGACTCC 1317
QY 1321 CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380
Db 1321 CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1377
QY 1381 TCTCAGGACAAGTGTCAACAGATGGAGAGGAAATGACCGGTTACATCGGAGAGTGTCA 1440
Db 1381 TCTCAGGACAAGTGTCAACAGATGGAGAGGAAATGACCGGTTACATCGGAGAGTGTCA 1437
QY 1441 GAGGTGAGGCTGTGCTTAGTCTAGAGAGGAGTGGAGCTGAGGCTTCTGAGACTCAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTAGTCTAGAGAGGAGTGGAGCTGAGGCTTCTGAGACTCAGAGA 1497

QY 1501 TCCCTCTTGAGCAGGACCTTGTCTACCTCATCATCAGAAATGCAGTAGCTTAAAGCGAAGT 1560
Db 1498 TCCCTCTTGAGCAGGACCTTGTCTACCTCATCATCAGAAATGCAGTAGCTTAAAGCGAAGT 1557
QY 1561 TTGAGCAAGCACGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCAAGCTTCTCAT 1620
Db 1558 TTGAGCAAGCACGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCAAGCTTCTCAT 1617
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTACCAAGCTCAA 1680
Db 1618 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTACCAAGCTCAA 1677
QY 1681 GTGAAAGAAATAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740
Db 1678 GTGAAAGAAATAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1737
QY 1741 CGGATGATCTCTACGAATCTGAGCTGAGAGGTCGCGCTTGTGCTGAAGAAATTCAG 1800
Db 1738 CGGATGATCTCTACGAATCTGAGCTGAGAGGTCGCGCTTGTGCTGAAGAAATTCAG 1797
QY 1801 CGGAAAGCGACAGAAATGTCAAGATAAACTGTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1860
Db 1798 CGGAAAGCGACAGAAATGTCAAGATAAACTGTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1857
QY 1861 GTGGAGAAATATGCGAAACTGGAGAAATCAATCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1858 GTGGAGAAATATGCGAAACTGGAGAAATCAATCTGAGCAGCAGCTCAAAATTCAGAG 1917
QY 1921 CTCGAAAGAAACTGGAGAAGCT----- 1944
Db 1918 CTCGAAAGAAACTGGAGAAGCTGTAAAGCAGCAGGAGGCCACCGAGCTGCTGCAG 1977
QY 1945 -----GCAAGGCGAGCGAGGAGGAGCTGGAGAAGCTTGAGAACTTGAGAACCCGAGAG 1992
Db 1978 AATATCCGCCAGGCAAGGAGCGAGCGAGGAGGAGCTGGAGAAGCTGCGAAGCCGAGAG 2037
QY 1993 GATTCTTTCTGAAGCATCAGAAAGACTGGTGAAGCTGAGGAACCGCGCATTTCTCTG 2052
Db 2038 GATTCTTTCTGAAGCATCAGAAAGACTGGTGAAGCTGAGGAACCGCGCATTTCTCTG 2097
QY 2053 GAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGGATGAC 2112
Db 2098 GAGAACAGGTAAAGAGACTAGAGACCATGAGCATGAGCGTAGAGAAAACAGACTGAAGGATGAC 2157
QY 2113 ATCAGACAAAATCCCAACAGATCCAGCAGATGCTGATATAAATTCCTGGAGCTCGAAGAG 2172
Db 2158 ATCAGACAAAATCCCAACAGATCCAGCAGATGCTGATATAAATTCCTGGAGCTCGAAGAG 2217
QY 2173 AAAATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAGTGCACTGAAAACAGAAAAGAG 2232
Db 2218 AAAATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAGTGCACTGAAAACAGAAAAGAG 2277
QY 2233 CAGCACTATGAGGAAAAGATTAAAGTGTGGAATGGAATCAAGATAAAGAAAGACCTGGCTGAC 2292
Db 2278 CAGCACTATGAGGAAAAGATTAAAGTGTGGAATGGAATCAAGATAAAGAAAGACCTGGCTGAC 2337
QY 2293 AAGGAGACACTGGGAGACATGATGAGAGACAGGAGGAGGCCCATGAGAAGGGCAAA 2352
Db 2338 AAGGAGACACTGGGAGACATGATGAGAGACAGGAGGAGGCCCATGAGAAGGGCAAA 2397
QY 2353 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGGAATCCAAAGATCAGATCCCTGGA 2412
Db 2398 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGGAATCCAAAGATCAGATCCCTGGA 2457
QY 2413 CAGGAGTTTGGAACTGTCTGAAGCCAAATAAATTCGACCAATAGCAGCTTTTATACC 2472
Db 2458 CAGGAGTTTGGAACTGTCTGAAGCCAAATAAATTCGACCAATAGCAGCTTTTATACC 2517
QY 2473 CAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2532
Db 2518 CAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2577

Db 4693 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTACCGCCCTTA 4752
Qy 4636 -----GAAAGAGAGAGCTGATGCTAAACTG 4662
Db 4753 GAATCAGTTGTGCGAGGTGGGAGTTTCTAGGGAAGAGCAGAGCTGATGCTAAACTG 4812
Qy 4663 CTTGGAACCTCCCTGCTGTAAGCTGGAAGTGATGACCGTCTAGACATGAACCTGACGCGTG 4722
Db 4813 CTTGGAACCTCCCTGCTGTAAGCTGGAAGTGATGACCGTCTAGACATGAACCTGACGCGTG 4872
Qy 4723 CCCTTCAGTGACCAAGTGGTGTGTGGGCAACGAGGAAGGGCTCTACGCCCTGAATGTC 4782
Db 4873 CCCTTCAGTGACCAAGTGGTGTGTGGGCAACGAGGAAGGGCTCTACGCCCTGAATGTC 4932
Qy 4783 TTGGAACCTCCCTAACCTGACCTCCAGAAATGAGGAGCTTCCAAATTTATATTATC 4842
Db 4933 TTGGAACCTCCCTAACCTGACCTCCAGAAATGAGGAGCTTCCAAATTTATATTATC 4992
Qy 4843 AAGGACCTCGAAGAGCTACTCATGTATGACGAGGAAGAGCGGCACTGTGCTTGTGGAC 4902
Db 4993 AAGGACCTCGAAGAGCTACTCATGTATGACGAGTGAAGAGCGGCACTGTGCTTGTGGAC 5052
Qy 4903 GTGAGAAAGTGAACAGTCCCTGGGCCAGTCCACCTGCTGCCAGCGGAGACATCTCA 4962
Db 5053 GTGAGAAAGTGAACAGTCCCTGGGCCAGTCCACCTGCTGCCAGCGGAGACATCTCA 5112
Qy 4963 CCCAACATTTTGAAGCTGTCAGGGCTGCCACTTGTGTGGGCGAGCAAGATTGAGAAC 5022
Db 5113 CCCAACATTTTGAAGCTGTCAGGGCTGCCACTTGTGTGGGCGAGCAAGATTGAGAAC 5172
Qy 5023 GGGCTCTGCATCTGTGACGCCATGCCCAGCAAGTGTGTCATTCCTCGGTACAAACGAAAC 5082
Db 5173 GGGCTCTGCATCTGTGACGCCATGCCCAGCAAGTGTGTCATTCCTCGGTACAAACGAAAC 5232
Qy 5083 CTCAGCAATATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAAGCTGTATCCAC 5142
Db 5233 CTCAGCAATATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAAGCTGTATCCAC 5292
Qy 5143 TTCACCAATTTACAGTATCTCTCATTTGAAACCAATAATTTACGAAATCGACATGAAGCAG 5202
Db 5293 TTCACCAATTTACAGTATCTCTCATTTGAAACCAATAATTTACGAAATCGACATGAAGCAG 5352
Qy 5203 TACAGCTTGAGGAATCTCTGGATAGATGACCATTCCTTGGCAGCTGCTGTTGGC 5262
Db 5353 TACAGCTTGAGGAATCTCTGGATAGATGACCATTCCTTGGCAGCTGCTGTTGGC 5412
Qy 5263 GCCTCTTCCAAAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGAGGCGAGCGAGAG 5322
Db 5413 GCCTCTTCCAAAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGAGGCGAGAGAG 5472
Qy 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGAATTTTACGGGAAGAGCTAGC 5382
Db 5473 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGAATTTTACGGGAAGAGCTAGC 5532
Qy 5383 CGCAGACGATCTCAAGTGGTGTGCTTACCTTTGGCCCTTTCCTACAGAGACCTAT 5442
Db 5533 CGCAGACGATCTCAAGTGGTGTGCTTACCTTTGGCCCTTTCCTACAGAGACCTAT 5592
Qy 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTAGATCCAGGACGCTCCTCAGCA 5502
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTAGATCCAGGACGCTCCTCAGCA 5652
Qy 5503 GGGACCCCTCCGAGGATCTGAGATCCCGAACCGCGCTTACCTGGCCCTTGCATT 5562
Db 5653 GGGACCCCTCCGAGGATCTGAGATCCCGAACCGCGCTTACCTGGCCCTTGCATT 5712
Qy 5563 TCCTCAGGAGGATTTACTTGGGCTCCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5622
Db 5713 TCCTCAGGAGGATTTACTTGGGCTCCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5772
Qy 5623 AAGGAAACCTGTGAGAGAGTCCGGCACTGAAACACCGGGGCGGCTCCACCTCCCGC 5682
Db 5773 AAGGAAACCTGTGAGAGAGTCCGGCACTGAAACACCGGGGCGGCTCCACCTCCCGC 5832

Qy 5683 AGCAGCCCCAACAGCGAGGCCCAACCGTACAAAGAGCACATCACCAAGCGCGTGGCC 5742
Db 5833 AGCAGCCCCAACAGCGAGGCCCAACCGTACAAAGAGCACATCACCAAGCGCGTGGCC 5892
Qy 5743 TCCAGCCCCAGCGCGCGCGAAGCGCCAGCCACCCGCGAGAGCCAGCACACCCACCGC 5802
Db 5893 TCCAGCCCCAGCGCGCGCGAAGCGCCAGCCACCCGCGAGAGCCAGCACACCCACCGC 5952
Qy 5803 TACCAGCGGGGCGGAGCCGAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCTTGGAGCGA 5862
Db 5953 TACCAGCGGGGCGGAGCCGAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCTTGGAGCGA 6012
Qy 5863 GAGAAGTCTCCCGCGCGCGATCTCAGCACCGGAGAGCGGTCTCCCGGGAGCGCTTTT 5922
Db 6013 GAGAAGTCTCCCGCGCGCGATCTCAGCACCGGAGAGCGGTCTCCCGGGAGCGCTTTT 6072
Qy 5923 GAAGACAGCAGCAGGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 5982
Db 6073 GAAGACAGCAGCAGGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6132
Qy 5983 AACAGGGAAGAGGGCA 5999
Db 6133 AACAGGTGAGGCAGCA 6149

RESULT 12
ABQ78871
ID ABQ78871 standard; cDNA; 5877 BP.
XX
AC ABQ78871;
XX
DT 10-OCT-2002 (first entry)
XX
DE Human kinase cDNA #2.
XX
KW Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;
KW Citron rho-interacting kinase; gene therapy; mental disorder; cancer;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..5877
FT /*tag= a
FT /product= "Kinase"
XX
PN WO200259325-A2.
XX
PD 01-AUG-2002.
XX
PF 20-DEC-2001; 2001WO-US050497.
XX
PR 27-DEC-2000; 2000US-0258335P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Yu X, Miranda M, Fridele CJ;
XX
DR WPI; 2002-599796/64.
DR P-PSDB; ABB81928.
XX
PT Novel polynucleotide encoding human proteins that are structurally
PT similar to animal kinases, useful for drug screening, diagnosis, in gene
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
PT applications.
XX
PS Disclosure; Page 44-45; 50pp; English.
XX
CC The invention relates to a novel human protein that shares structural
CC similarity with animal kinases, including serine-threonine kinases,
CC particularly Citron rho-interacting kinases. The proteins of the
CC invention have nootropic and cytostatic activity. The polynucleotides may

CC have a use in gene therapy. The encoded novel polypeptides are useful for
CC generating antibodies, as reagents in diagnostic assays, for identifying
CC other cellular gene products related to NHP and as reagents in assays for
CC screening for compounds that are useful in the treatment of mental,
CC biological or medical disorders and diseases including cancer. The
CC sequence encodes a novel human kinase of the invention
XX
SQ

Sequence 5877 BP; 1680 A; 1433 C; 1583 G; 1181 T; 0 U; 0 Other;

Query Match 87.0%; Score 5358; DB 6; Length 5877;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5634; Conservative 0; Mismatches 5; Indels 243; Gaps 3;

```
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTGGATGCTGGTCTGCTGAACCCATTT 60
DB |||||
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTGGATGCTGGTCTGCTGAACCCATTT 60
DB |||||
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA 120
DB |||||
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA 120
DB |||||
QY 121 CAGCAGATGCTCCTCTTCCGAGAGGGAATATTAGATGCCCTCTTTGTCCTTTTGA 180
DB |||||
QY 121 CAGCAGATGCTCCTCTTCCGAGAGGGAATATTAGATGCCCTCTTTGTCCTTTTGA 180
DB |||||
QY 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCAGTGAGCAACTTTGTCGGGAAGTAT 240
DB |||||
QY 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAAGCAGTGAGCAACTTTGTCGGGAAGTAT 240
DB |||||
QY 241 TCGGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGGACTTCGAAGTCAGA 300
DB |||||
QY 241 TCGGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGGACTTCGAAGTCAGA 300
DB |||||
QY 301 AGTCTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB |||||
QY 301 AGTCTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB |||||
QY 361 GACATCTATCTATGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB |||||
QY 361 GACATCTATCTATGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB |||||
QY 421 TTTTGTGAGAGAGCGGACATATTATCTGAGGACACAGCCCGTGGATCCGCCAATTA 480
DB |||||
QY 421 TTTTGTGAGAGAGCGGACATATTATCTGAGGACACAGCCCGTGGATCCGCCAATTA 480
DB |||||
QY 481 CAGTATGCTCTTCAGGACAAATATCACTTTATCTGATGGAGAAATATCAGCTTGAGGG 540
DB |||||
QY 481 CAGTATGCTCTTCAGGACAAATATCACTTTATCTGATGGAGAAATATCAGCTTGAGGG 540
DB |||||
QY 541 GACTTGGTGTCACTTTGATAGATATGAGGACAGTATGATGAAATCTGATACAGTTT 600
DB |||||
QY 541 GACTTGGTGTCACTTTGATAGATATGAGGACAGTATGATGAAATCTGATACAGTTT 600
DB |||||
QY 601 TACTAGCTGAGCTGATTTTGGCTGTTTACAGGTTTCTATCTGATGGGATACCTGATCGA 660
DB |||||
QY 601 TACTAGCTGAGCTGATTTTGGCTGTTTACAGGTTTCTATCTGATGGGATACCTGATCGA 660
DB |||||
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
DB |||||
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
DB |||||
QY 721 GGAATCTGCCGGAATGAATCAACAAGATGGTGAATGCCAACTCCCGATGGGACC 780
DB |||||
QY 721 GGAATCTGCCGGAATGAATCAACAAGATGGTGAATGCCAACTCCCGATGGGACC 780
DB |||||
QY 781 CCAGATTACATGGCTCCTGAAGTGTGATGATGACAGGGGATCGAAAGGCCACTTAC 840
DB |||||
QY 781 CCAGATTACATGGCTCCTGAAGTGTGATGATGACAGGGGATCGAAAGGCCACTTAC 840
DB |||||
QY 841 GGCCTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB |||||
QY 841 GGCCTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB |||||
```

```
QY 901 TCCCCCTTGCAGAGGAAACCTCTCCAGAACTTCAATAACATTATGAATTTCCAGCGG 960
DB |||||
QY 901 TCCCCCTTGCAGAGGAAACCTCTCCAGAACTTCAATAACATTATGAATTTCCAGCGG 960
DB |||||
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGGAGTGAATTTCTTGTATCTGATCTCAAAAGC 1020
DB |||||
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGGAGTGAATTTCTTGTATCTGATCTCAAAAGC 1020
DB |||||
QY 1021 TTGTTGTGCGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGGTGCATCTCTTTCTTC 1080
DB |||||
QY 1021 TTGTTGTGCGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGGTGCATCTCTTTCTTC 1080
DB |||||
QY 1081 TCTAAATTTGATGTGAACCAACATTTCTGTAACCTCTCTCCCTCTTCTGTTCCCACTCAAG 1140
DB |||||
QY 1081 TCTAAATTTGATGTGAACCAACATTTCTGTAACCTCTCTCCCTCTTCTGTTCCCACTCAAG 1140
DB |||||
QY 1141 TCTGACGATGACCTCCAAATTTTGTGATGAACAGAGAGAAATTTGTTGGGTTTCTCTCT 1200
DB |||||
QY 1141 TCTGACGATGACCTCCAAATTTTGTGATGAACAGAGAGAAATTTGTTGGGTTTCTCTCT 1200
DB |||||
QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGTTTTCG 1260
DB |||||
QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGTTTTCG 1260
DB |||||
QY 1261 TACGCAAGGCACTGGGATTTCTGTTAGATCTGAGTCTGTTGTGGGTTTCTGACTCC 1320
DB |||||
QY 1261 TACGCAAGGCACTGGGATTTCTGTTAGATCTGAGTCTGTTGTGGGTTTCTGACTCC 1320
DB |||||
QY 1321 CCTGCCAAGCTAGCTCCATCGAAAGAACTTCTCTCAATAAAGCAAGAGCTACAAGAC 1380
DB |||||
QY 1321 CCTGCCAAGCTAGCTCCATCGAAAGAACTTCTCTCAATAAAGCAAGAGCTACAAGAC 1380
DB |||||
QY 1381 TCTCAGGCAAGTGTCAAGATGCGAGCAAGAAATGACCGGTTTACATCGGAGGTTCTCA 1440
DB |||||
QY 1381 TCTCAGGCAAGTGTCAAGATGCGAGCAAGAAATGACCGGTTTACATCGGAGGTTCTCA 1440
DB |||||
QY 1441 GAGTGGAGGCTGTGCTTAGTTCAGAGAGGTTGAGGCTGAAGGCTCTGAGACTCAGAGA 1500
DB |||||
QY 1441 GAGTGGAGGCTGTGCTTAGTTCAGAGAGGTTGAGGCTGAAGGCTCTGAGACTCAGAGA 1500
DB |||||
QY 1501 TCCCTCTCGAGCAGGACCTTGTACCTACATACAGAAATGCGAGTAGCTTAAAGCGAAGT 1560
DB |||||
QY 1501 TCCCTCTCGAGCAGGACCTTGTACCTACATACAGAAATGCGAGTAGCTTAAAGCGAAGT 1560
DB |||||
QY 1561 TTGAGCAAGCAGCGATGGAGGTGTTCCAGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1620
DB |||||
QY 1561 TTGAGCAAGCAGCGATGGAGGTGTTCCAGAGGATGACAAAGCACTGCGAGCTTCTCCAT 1620
DB |||||
QY 1621 GATATCAGAGCAGAGCAGCGGAGCTCCAAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA 1680
DB |||||
QY 1621 GATATCAGAGCAGAGCAGCGGAGCTCCAAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA 1680
DB |||||
QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB |||||
QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB |||||
QY 1741 CGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800
DB |||||
QY 1741 CGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800
DB |||||
QY 1801 CGGAAAGCAGAGTCTCAGCAATAAATGTTTGAAGCTTAAGGATCAAGGAGGCTGAA 1860
DB |||||
QY 1801 CGGAAAGCAGAGTCTCAGCAATAAATGTTTGAAGCTTAAGGATCAAGGAGGCTGAA 1860
DB |||||
QY 1861 GTGGGAGAAATTCGAAAATGGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTTCAAGG 1920
DB |||||
QY 1861 GTGGGAGAAATTCGAAAATGGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTTCAAGG 1920
DB |||||
QY 1921 CTCCAGAGAAATCTGAGAAAGGCT----- 1944
DB |||||
QY 1921 CTCCAGAGAAATCTGAGAAAGGCTTGAAAAGCCAGCAGGAGCCACCGAGCTGTGAG 1980
DB |||||
QY 1945 -----GCAAAAGGAGCGAGCGGAGCTGGAGAGCTGGAGAAAGCTGCAGAACCGAGAG 1992
DB |||||
```

Db 1981 AATATCCGCCGCAAGAGCGAGCGGAGCTGGAGAGCTCGAGAACCGAG 2040
Qy 1993 GATTTCTTCTGAAGGATCAGAAAGAGCTGTGGAGCTGAGGAACGCCGCTTCTG 2052
Db 2041 GATTTCTTCTGAAGGATCAGAAAGAGCTGTGGAGCTGAGGAACGCCGCTTCTG 2100
Qy 2053 GAGAACAGGTAAAGAGACTAGAGCCATGGAGCTAGAGAAACAGACTGAGGATGAC 2112
Db 2101 GAGAACAGGTAAAGAGACTAGAGCCATGGAGCTAGAGAAACAGACTGAGGATGAC 2160
Qy 2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAG 2172
Db 2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAG 2220
Qy 2173 AAAATCGGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAGTCCACTGAAACAGAAAG 2232
Db 2221 AAAATCGGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAGTCCACTGAAACAGAAAG 2280
Qy 2233 CAGCACTATGAGGAAAAGATTAAGTGTGGACAATCAGATAAAGAAAGCTGCTGAC 2292
Db 2281 CAGCACTATGAGGAAAAGATTAAGTGTGGACAATCAGATAAAGAAAGCTGCTGAC 2340
Qy 2293 AAGGAGACTCGGAAACATGATGAGAGACACGAGGAGGCGCCATGAGAAAGGCAAA 2352
Db 2341 AAGGAGACTCGGAAACATGATGAGAGACACGAGGAGGCGCCATGAGAAAGGCAAA 2400
Qy 2353 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGATGATCCAGATCAGATCCCTGGAA 2412
Db 2401 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGATGATCCAGATCAGATCCCTGGAA 2460
Qy 2413 CAGAGATTGTGGAACTGTCTGAAGCCAAATAAATTTGACGAAATAGCAGTCTTTTACC 2472
Db 2461 CAGAGATTGTGGAACTGTCTGAAGCCAAATAAATTTGACGAAATAGCAGTCTTTTACC 2520
Qy 2473 CAAAGGAAATGAAGGCCCAAGAGAGATGATTTCTGAATCAGGCAACAGAGAAATTTAC 2532
Db 2521 CAAAGGAAATGAAGGCCCAAGAGAGATGATTTCTGAATCAGGCAACAGAGAAATTTAC 2580
Qy 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAAATCTGGAGGAGCAGCTGGAG 2592
Db 2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAAATCTGGAGGAGCAGCTGGAG 2640
Qy 2593 AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTCTGGAACTGGAGACAAGATTG 2652
Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTCTGGAACTGGAGACAAGATTG 2700
Qy 2653 CGGGAGTCACTAGAGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAGAG 2712
Db 2701 CGGGAGTCACTAGAGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAGAG 2760
Qy 2713 CTACAGCTCTCCCTGAGGAGCGCGAGTCACAGTTGACAGCCCTGACGGCTGCACGGCG 2772
Db 2761 CTACAGCTCTCCCTGAGGAGCGCGAGTCACAGTTGACAGCCCTGACGGCTGCACGGCG 2820
Qy 2773 GCCTGGAGAGCCAGCTTGGCAGGCGAAGACAGAGCTGGAAGAGACACAGCAGAGCT 2832
Db 2821 GCCTGGAGAGCCAGCTTGGCAGGCGAAGACAGAGCTGGAAGAGACACAGCAGAGCT 2880
Qy 2833 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2892
Db 2881 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940
Qy 2893 CTTTCGTAACAGCTGTACTGTAATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952
Db 2941 CTTTCGTAACAGCTGTACTGTAATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 3000
Qy 2953 GACACGCTGAACTCAACACCAAAATTTCTGCTGCAAACTCGATGAGGCTTCT 3012
Db 3001 GACACGCTGAACTCAACACCAAAATTTCTGCTGCAAACTCGATGAGGCTTCT 3060
Qy 3013 GCGGCCAACGAGAGATTGTACAATTCGGAAGTGAAGCCATCTCCGCCGCGAGATC 3072

Db 3061 GGCGCCAACGACGAGATTGTACAACCTGCGAAGTGAAGTGACCATCTCCGCCGCGAGATC 3120
Qy 3073 ACGGAACGAGAGATGAGCTTACAGCCAGAAACGAAACGATGAGGCTCTGAGAACCAAG 3132
Db 3121 ACGGAACGAGAGATGAGCTTACAGCCAGAAACGAAACGATGAGGCTCTGAGAACCAAG 3180
Qy 3133 TGCACCATGCTGGAGGAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3192
Db 3181 TGCACCATGCTGGAGGAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3240
Qy 3193 AAAGAGCGGAGTGGAGGCGCTGGAGAGGCTGCTGGGTGATGAGAAATCCAGTTTGGAG 3252
Db 3241 AAAGAGCGGAGTGGAGGCGCTGGAGAGGCTGCTGGGTGATGAGAAATCCAGTTTGGAG 3300
Qy 3253 TGTCCGGTTCGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGACGGGCGAGAGCC 3312
Db 3301 TGTCCGGTTCGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGACGGGCGAGAGCC 3360
Qy 3313 GATCAGCGGATCACCGAGTCTCCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT 3372
Db 3361 GATCAGCGGATCACCGAGTCTCCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT 3420
Qy 3373 GAGATTTCTGCTCTGACGAGGCTCTCAAGAGACAGAGCTGAGGCGCGAGAGCTCTCT 3432
Db 3421 GAGATTTCTGCTCTGACGAGGCTCTCAAGAGACAGAGCTGAGGCGCGAGAGCTCTCT 3480
Qy 3433 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAATGAAATGCGCCGAGCTTA 3492
Db 3481 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAATGAAATGCGCCGAGCTTA 3540
Qy 3493 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCCAAA 3552
Db 3541 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCCAAA 3600
Qy 3553 TTACAGCAGCAGATGGAGCTGCAAAAAATCACATTTTCCGCTCTGACTCAAGGACTCACA 3612
Db 3601 TTACAGCAGCAGATGGAGCTGCAAAAAATCACATTTTCCGCTCTGACTCAAGGACTCACA 3660
Qy 3613 GAACTCTAGATCGGGTGTATCTACTGAAGACAGAAAGATGACTTGGAGATATCAGCTG 3672
Db 3661 GAACTCTAGATCGGGTGTATCTACTGAAGACAGAAAGATGACTTGGAGATATCAGCTG 3720
Qy 3673 GAAACATTCAGGTTCTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3732
Db 3721 GAAACATTCAGGTTCTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3780
Qy 3733 CAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGACCAACCTGCTAAAAAGAAAAG 3792
Db 3781 CAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGACCAACCTGCTAAAAAGAAAAG -- 3838
Qy 3793 GGTTTATTAGTCAGCGAAAGAGGACCTGCTTATCCCAACAGGTTCTCTGCGATAC 3852
Db 3839 -----AGGTTCTCTGCGATAC 3855
Qy 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGGCC 3912
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGGCC 3915
Qy 3913 CTTCAGAAGACCCGATCGAGCTCCGCTCCGCCGCGGAGGAGCTGCCACCGCAAGAGCA 3972
Db 3916 CTTCAGAAGACCCGATCGAGCTCCGCTCCGCCGCGGAGGAGCTGCCACCGCAAGAGCA 3975
Qy 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGCGAGGAGCAGAGATCGCCATCTCCGCC 4032
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGCGAGGAGCAGAGATCGCCATCTCCGCC 4035
Qy 4033 ATCTGCGGTGCGCAGAGCAGCCAGCCAGTGCCTATGAGCTGTGGCCCGCCCATCCAGC 4092
Db 4036 ATCTGCGGTGCGCAGAGCAGCCAGTGCCTATGAGCTGTGGCCCGCCCATCCAGC 4095
Qy 4093 CGCAGAAAGGAGTCTTCACTCCAGAGAAATTTAGTGGGCTCTTAAGGAGCAGATCCAC 4152
Db 4096 CGCAGAAAGGAGTCTTCACTCCAGAGAAATTTAGTGGGCTCTTAAGGAGCAGATCCAC 4155

```
QY 4153 CACAAATATTCCTCACCGATTCAAGTAGGACTGAAATGCGAGCCACAAAGTGCTGTG 4212
Db 4156 CACAATATTCCTCACCGATTCAAGTAGGACTGAAATGCGAGCCACAAAGTGCTGTG 4215
QY 4213 TGTCTGATACCGTGACATTTGGACGCGCGGCATCCAAATGCTCCGAATGTCAGGTGATG 4272
Db 4216 TGTCTGATACCGTGACATTTGGACGCGCGGCATCCAAATGCTCCGAATGTCAGGTGATG 4275
QY 4273 TGTACACCCCAAGTGTCTCCAGTGTGTCAGCAGCACTGCGGCTGTGCTGTAATGCC 4332
Db 4276 TGTACACCCCAAGTGTCTCCAGTGTGTCAGCAGCACTGCGGCTGTGCTGTAATGCC 4335
QY 4333 ACACATTCACCGAGGCTTCTCCGCTGACAAATGAATCCCCAGGTCTCCAGACCAAG 4392
Db 4336 ACACATTCACCGAGGCTTCTCCGCTGACAAATGAATCCCCAGGTCTCCAGACCAAG 4395
QY 4393 GAGCCAGCAGAGCTTGCACTTGGAAAGGTGATGAAGTGCAGGAATACCAACGA 4452
Db 4396 GAGCCAGCAGAGCTTGCACTTGGAAAGGTGATGAAGTGCAGGAATACCAACGA 4455
QY 4453 GGACAGCAAGCTGGACAGAGTACATTTGCTCGAGGGATCAAAAGTCTCATTTAT 4512
Db 4456 GGACAGCAAGCTGGACAGAGTACATTTGCTCGAGGGATCAAAAGTCTCATTTAT 4515
QY 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAAAGATTTGAGCTGTGCCCTTCC 4572
Db 4516 GACAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAAAGATTTGAGCTGTGCCCTTCC 4575
QY 4573 GAGGGGATGTATCTATTCATTTGCTGCGCTTGGTCTTCGAACTCGCAATACAGCCAAA 4632
Db 4576 GAGGGGATGTATCTATTCATTTGCTGCGCTTGGTCTTCGAACTCGCAATACAGCCAAA 4635
QY 4633 GCA----- 4635
Db 4636 GCAGATGTCCATACATCTGAGATGGAATCTCACCCGACACACCTGCTGGCCCGG 4695
QY 4636 ----- 4635
Db 4696 AGAACCTCTACTTGCTAGCTCCAGCTTCCCTGTGACAAACAGCGTGGGTCAACCGCTTA 4755
QY 4636 -----GAAAGCAGAGCTGATGCTTAACTG 4662
Db 4756 GAATCAGTTGTCGACGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTTAACTG 4815
QY 4663 CTTGGAATCTCCTGCTGAACTGGAAGTGTATGACCGTCTAGACATGAATGCACGCTG 4722
Db 4816 CTTGGAATCTCCTGCTGAACTGGAAGTGTATGACCGTCTAGACATGAATGCACGCTG 4875
QY 4723 CCGTTCAGTGACAGGTGTGTTGGTGGCAGCGAGAGGGCTCTACGCCCTGATGTC 4782
Db 4876 CCGTTCAGTGACAGGTGTGTTGGTGGCAGCGAGAGGGCTCTACGCCCTGATGTC 4935
QY 4783 TTGAAAACTCCCTAACCCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4842
Db 4936 TTGAAAACTCCCTAACCCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4995
QY 4843 AAGGACCTGGAGAGTACTCATGTATAGCAGAGAGAGCGGCACTGTGCTTTGTGAC 4902
Db 4996 AAGGACCTGGAGAGTACTCATGTATAGCAGAGAGAGCGGCACTGTGCTTTGTGAC 5055
QY 4903 GTCAAGAAATGTAACAGTCCCTGGCCAGTCCCACTGCTCCGTCAGCCAGCGACATCTCA 4962
Db 5056 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTCCGTCAGCCAGCGACATCTCA 5115
QY 4963 CCCAATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGCGAGGCAAGATTGAGAAC 5022
Db 5116 CCCAATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGCGAGGCAAGATTGAGAAC 5175
QY 5023 GGGCTCTGCTATCTGTCAGCCATGCCAGCAAGTCTCATTTCTCCGCTACAAAGAAAC 5082
Db 5176 GGGCTCTGCTATCTGTCAGCCATGCCAGCAAGTCTCATTTCTCCGCTACAAAGAAAC 5235
```

```
QY 5083 CTCACCAATTAATCGATCCGGAAGAGATAGACCTCAGACCTCAGAGCCCTGAGCTGTATCCAC 5142
Db 5236 CTCACCAATTAATCGATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
QY 5143 TTCACCAATTAATCGATCCCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5202
Db 5296 TTCACCAATTAATCGATCCCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
QY 5203 TACACGCTCGAGGAATTCCTCGATAGAATGAACATTTCTTTGGCACTGTGTGTTGCC 5262
Db 5356 TACACGCTCGAGGAATTCCTCGATAGAATGAACATTTCTTTGGCACTGTGTGTTGCC 5415
QY 5263 GCCTCTTCCAAACAGTTCCTCTGCTCAATCGTCAGGTGAACAGCGAGGCGCAGGAG 5322
Db 5416 GCCTCTTCCAAACAGTTCCTCTGCTCAATCGTCAGGTGAACAGCGAGGCGCAGGAG 5475
QY 5323 GAGTACTTGTGCTGTTTCCACGAAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC 5382
Db 5476 GAGTACTTGTGCTGTTTCCACGAAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC 5535
QY 5383 CGCACAGAGGATCTCAAGTGGAGTGCCTTACCTTTGGCTTTGCTTACAGAAACCTAT 5442
Db 5536 CGCACAGAGGATCTCAAGTGGAGTGCCTTACCTTTGGCTTTGCTTACAGAAACCTAT 5595
QY 5443 CTGTTTGTGACCCCACTTCACTCACTCGAAATTAATGAGATCCAGGACCGCTCCTCAGCA 5502
Db 5596 CTGTTTGTGACCCCACTTCACTCACTCGAAATTAATGAGATCCAGGACCGCTCCTCAGCA 5655
QY 5503 GGGACCCCTGCCGAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5562
Db 5656 GGGACCCCTGCCGAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5715
QY 5563 TCCTCAGAGCGATTTACTTTGGGCTCCTCATACAGGATAAATTAAGGGTCAATTGCTGC 5622
Db 5716 TCCTCAGAGCGATTTACTTTGGGCTCCTCATACAGGATAAATTAAGGGTCAATTGCTGC 5775
QY 5623 AAGGAAACCTCGTGAAGAGTCCGGCACTGAACACACCCGGGGCGCTCCACTCCCGC 5682
Db 5776 AAGGAAACCTCGTGAAGAGTCCGGCACTGAACACACCCGGGGCGCTCCACTCCCGC 5835
QY 5683 AG 5684
Db 5836 AG 5837

RESULT 13
AAC77568
ID AAC77568 standard; cDNA; 6609 BP.
XX
AC AAC77568;
XX
DT 08-FEB-2001 (first entry)
XX
Human OREF3123 polynucleotide sequence SEQ ID NO:6245.
XX
Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antiinflammatory;
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
FN W020058473-A2.
```

XX 05-OCT-2000.
PD 31-MAR-2000; 2000WO-US008621.
XX 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
XX WPI; 2000-602362/57.
DR P-PSDB; AAB43359.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX Claim 5; Page 5429-5433; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX enhance coagulation; to inhibit thrombosis; and as a contraceptive
SQ Sequence 6609 BP; 1694 A; 1779 C; 1646 G; 1490 T; 0 U; 0 Other;
Query Match 56.4%; Score 3475.2; DB 3; Length 6609;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 3710; Conservative 0; Mismatches 33; Indels 195; Gaps 2;
QY 2255 AAGTGTGGACATCAGATAAAGAAAGACCTGGCTGACAGAGAGACACTGGAGAACATGA 2314
DB 299 AGGTGTTGGACATCAGATAAAGAAAGACCTGGCTGACAGAGAGACACTGGAGAACATGA 358
QY 2315 TGACAGACACAGAGAGAGAGGCGCCATGAGAGGCGAATTTCTACGCGAACAGAGGGCA 2374
DB 359 TGACAGACACAGAGAGAGAGGCGCCATGAGAGGCGAATTTCTACGCGAACAGAGGGCA 418
QY 2375 TGATCAATGCTATGGATTCCAAAGATCAGATCCCTCGGAACAGAGGATTGTGGAACCTGCTG 2434
DB 419 TGATCAATGCTATGGATTCCAAAGATCAGATCCCTCGGAACAGAGGATTGTGGAACCTGCTG 478
QY 2435 AAGCCAAATAACTTGGAGCAATAGACATCTTTTACCCAAAGGAACATGAAGGCCCAAG 2494
DB 479 AAGCCAAATAACTTGGAGCAATAGACATCTTTTACCCAAAGGAACATGAAGGCCCAAG 538
QY 2495 AAGAGATGATTTCTGAACTCAGGCAACAGAAAATTTTACCTGGAGACACAGGCTGGGAAGT 2554
DB 539 AAGAGATGATTTCTGAACTCAGGCAACAGAAAATTTTACCTGGAGACACAGGCTGGGAAGT 598
QY 2555 TGGAGGCCCAAGAACCGAAAATCTGGAGAGACAGCTGGAGAGATTCAGCCACCAAGACCACA 2614
DB 599 TGGAGGCCCAAGAACCGAAAATCTGGAGAGACAGCTGGAGAGATTCAGCCACCAAGACCACA 658

QY 2615 GTGACAGAAATCGGCTCTGAACTCGAGACAGAGATTCCGGAGGTCAGTCTAGAGCAGC 2674
DB 659 GTGACAGAAATCGGCTCTGAACTCGAGACAGAGATTCCGGAGGTCAGTCTAGAGCAGC 718
QY 2675 AGGAGCAGAAATCGAGCTCAAGCGCCAGCTTCACAGAGCTACAGCTCTCCCTGCGAGGAGC 2734
DB 719 AGGAGCAGAAATCGAGCTCAAGCGCCAGCTTCACAGAGCTACAGCTCTCCCTGCGAGGAGC 778
QY 2735 GCGAGTCAAGTTGACAGCTCTGAGCTGCAAGCGCGCCCTGGAGAGCCAGCTTGGCC 2794
DB 779 GCGAGTCAAGTTGACAGCTCTGAGCTGCAAGCGCGCCCTGGAGAGCCAGCTTGGCC 838
QY 2795 AGGCGAAGACAGAGCTGGAAGAGACCAAGCAGAGCTGAAGAGAGATCCAGGCACTCA 2854
DB 839 AGGCGAAGACAGAGCTGGAAGAGACCAAGCAGAGCTGAAGAGAGATCCAGGCACTCA 898
QY 2855 CGGCACATAGAGATGAAATCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGACTGTAA 2914
DB 899 CGGCACATAGAGATGAAATCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGACTGTAA 958
QY 2915 TCACAGACCTGGAGGAGCAGTAAACAGCTGACCGAGGACACCGCTGAACCTCAACCAAC 2974
DB 959 TCACAGACCTGGAGGAGCAGTAAACAGCTGACCGAGGACACCGCTGAACCTCAACCAAC 1018
QY 2975 AAAAATCTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCCAACGACAGATTGTAC 3034
DB 1019 AAAAATCTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCCAACGACAGATTGTAC 1078
QY 3035 AACTGCGAAGTGAAGTGAACCTCTCGCGGAGATCACGGAAACGAGAGATGACGCTTA 3094
DB 1079 AACTGCGAAGTGAAGTGAACCTCTCGCGGAGATCACGGAAACGAGAGATGACGCTTA 1138
QY 3095 CCAGCCAGAAACAAACGATGAGGCTCTGAAGACACGCTGCACCATCTGGAGGAAACAGG 3154
DB 1139 CCAGCCAGAAACAAACGATGAGGCTCTGAAGACACGCTGCACCATCTGGAGGAAACAGG 1198
QY 3155 TCATGGATTTGGAGGCTTAAACGATGAGTGTCTAGAAAAGAGCGGAGTGGAGGCTC 3214
DB 1199 TCATGGATTTGGAGGCTTAAACGATGAGTGTCTAGAAAAGAGCGGAGTGGAGGCTC 1258
QY 3215 GGGAGGCGCTCTGGGTGATGAGAAATCCAGATTTGAGTGTCCGGTTTCGAGAGCTGAGA 3274
DB 1259 GGGAGGCGCTCTGGGTGATGAGAAATCCAGATTTGAGTGTCCGGTTTCGAGAGCTGAGA 1318
QY 3275 GAATGCTGGACACCGAGAAACAGACAGGCGGAGCGGATCAGCGGATCACCGAGTCTC 3334
DB 1319 GAATGCTGGACACCGAGAAACAGACAGGCGGAGCGGATCAGCGGATCACCGAGTCTC 1378
QY 3335 GCCAGGTGGAGCTGGAGTGAAGAGACACAGGCTGAGATTCCTGCTTCGAGCAGG 3394
DB 1379 GCCAGGTGGAGCTGGAGTGAAGAGACACAGGCTGAGATTCCTGCTTCGAGCAGG 1438
QY 3395 CTCTCAAGAGACAGAAAGCTGAAGCGCCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGA 3454
DB 1439 CTCTCAAGAGACAGAAAGCTGAAGCGCCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGA 1498
QY 3455 AGAAGCATGCTATCTGTAATGAATGCCGAGCTTACAGAGAGCTGAGAGCTGAGACTGAAC 3514
DB 1499 AGAAGCATGCTATCTGTAATGAATGCCGAGCTTACAGAGAGCTGAGAGCTGAGACTGAAC 1558
QY 3515 GAGAGCTCAACACAGAGCTTCTGGAAGCAAGCCAAATTTACAGCAGCAGATGGACCTGC 3574
DB 1559 GAGAGCTCAACACAGAGCTTCTGGAAGCAAGCCAAATTTACAGCAGCAGATGGACCTGC 1618
QY 3575 AGAAAAATCAATTTTCGCTGTGACTCAAGGACTTCGAAAGCTCTTAGATCGGGCTGATC 3634
DB 1619 AGAAAAATCAATTTTCGCTGTGACTCAAGGACTTCGAAAGCTCTTAGATCGGGCTGATC 1678
QY 3635 TACTGGAAGACAGAAAGAGTGAATTCGAGTATCAGCTGGAAACATTCAGGTTCTCTATT 3694
DB 1679 TACTGGAAGACAGAAAGAGTGAATTCGAGTATCAGCTGGAAACATTCAGGTTCTCTATT 1738

QY 3695 CTGATGAAGGTGAAGTGAAGGCACTATTCTCTCAACAAACCAAACTCATTTGATTTC 3754
Db 1739 CTGATGAAGGTGAAGTGAAGGCACTATTCTCTCAACAAACCAAACTCATTTGATTTC 1798
QY 3755 TGAAGGCAAAATGGAACCACTGCTGCTAAAAGAAAAGGGTTTATTAGTCGACGGAAG 3814
Db 1799 TGAAGGCAAAATGGAACCACTGCTGCTAAAAGAAAAG 1834
QY 3815 AGGACCTGCTTTACCAACACAGAGTTCCTCTGAGGTACAAATGAGCTGAAGCTGGCCCTGG 3874
Db 1835 -----AGGTTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGG 1873
QY 3875 AGAAGGAGAAAGTCTGCTGTGACAGCTAGAGAAAGCCCTTTCAGAAAGCCGATCGAGC 3934
Db 1874 AGAAGGAGAAAGTCTGCTGTGACAGCTAGAGAAAGCCCTTTCAGAAAGCCGATCGAGC 1933
QY 3935 TCCGGTCCGCGGAGGAGAGCTGCCCAACCGCAAGAAAGAAAGCAACCAACCCATCCA 3994
Db 1934 TCCGGTCCGCGGAGGAGAGCTGCCCAACCGCAAGAAAGCAACCAACCCATCCA 1993
QY 3995 CGCAGCCACCGGAGGAGAGCTGCCCAATGCGCATGTCGCGCATGTCGCGTGCAGAGCACC 4054
Db 1994 CGCAGCCACCGGAGGAGAGCTGCCCAATGCGCATGTCGCGCATGTCGCGTGCAGAGCACC 2053
QY 4055 AGCCAGTGCATGAGCCTGCTGCGCCCGCCATCCAGCCGCGAGAAAGAGTCTTCAATC 4114
Db 2054 AGCCAGTGCATGAGCCTGCTGCGCCCGCCATCCAGCCGCGAGAAAGAGTCTTCAATC 2113
QY 4115 CAGAGGAATTTAGTCGGCTTTAAGGAAGCATGACACCAATATTCCTACCGATTCA 4174
Db 2114 CAGAGGAATTTAGTCGGCTTTAAGGAAGCATGACACCAATATTCCTACCGATTCA 2173
QY 4175 ACGTAGACTGAACATCGGAGCACAAGGTGCTGTGTCTGGATACCGTGCACTTTG 4234
Db 2174 ACGTAGACTGAACATCGGAGCACAAGGTGCTGTGTCTGGATACCGTGCACTTTG 2233
QY 4235 GAGCCAGGCAATCCAAATGCTCGAATGTGAGTGTGATGTGTCACCCAAAGTGTCCACGT 4294
Db 2234 GAGCCAGGCAATCCAAATGCTCGAATGTGAGTGTGATGTGTCACCCAAAGTGTCCACGT 2293
QY 4295 GCTTGCAGCACAATCGGGCTGCTGCTGAATATGCCACACACTTACCGAGGCTTCT 4354
Db 2294 GCTTGCAGCACAATCGGGCTGCTGCTGAATATGCCACACACTTACCGAGGCTTCT 2353
QY 4355 GCCGTGACAAAATGAATCCCAAGTCTCCAGACCAAGAGGCCAGCAGCTTGCACC 4414
Db 2354 GCCGTGACAAAATGAATCCCAAGTCTCCAGACCAAGAGGCCAGCAGCTTGCACC 2413
QY 4415 TGGAGGTTGGATGAAGGTGCCAGGAATAACAAACGAGGACAGCAGGCTGGACAGGA 4474
Db 2414 TGGAGGTTGGATGAAGGTGCCAGGAATAACAAACGAGGACAGCAGGCTGGACAGGA 2473
QY 4475 AGTACATTTGCTGGAGGATCAAAAGTCTCTCATTTATGACAAATGAAGCCAGAGAGCTG 4534
Db 2474 AGTACATTTGCTGGAGGATCAAAAGTCTCTCATTTATGACAAATGAAGCCAGAGAGCTG 2533
QY 4535 GACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCCGAGGGGATGTATCTATTGATG 4594
Db 2534 GACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCCGAGGGGATGTATCTATTGATG 2593
QY 4595 GTCCGCTGGTGTCTCCGAACTCGCAATACAGCCAAAGCA----- 4635
Db 2594 GTCCGCTGGTGTCTCCGAACTCGCAATACAGCCAAAGCA----- 4635
QY 4636 ----- 4635
Db 2654 AGATGGAATCTCACCGGCACACCACTGCTGGCCCGGAGAACCTCTACTTGTAGTCT 2713
QY 4636 ----- 4635
Db 2714 CAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTGAATCAGTTGTCGAGGTGGGA 2773
QY 4636 -----GAAAAGCAGAGCTGATGCTAAACTGCTTGGAAACTCCCTGCTGAAAC 4684

Db 2774 GAGTTTCTAGGAAAAAACAGAGCTGATGCTAACTGCTTGGAAACTCCCTGCTGAAAC 2833
QY 4685 TGAAGGTGATGACCGCTTAGACATGAACCTGCAGCTGCCCTTCAGTGACAGAGTGTTG 4744
Db 2834 TGAAGGTGATGACCGCTTAGACATGAACCTGCAGCTGCCCTTCAGTGACAGAGTGTTG 2893
QY 4745 TGGTGGGCAACGAGAGAGGCTCTACGCCCTGAAATGCTTGAAGAACTCCCTAACCCATG 4804
Db 2894 TGGTGGGCAACGAGAGAGGCTCTACGCCCTGAAATGCTTGAAGAACTCCCTAACCCATG 2953
QY 4805 TCCAGGAATTTGAGAGCAGTCTTCCAAATTTATATCAAGGACCTGGAGAAGTACTCA 4864
Db 2954 TCCAGGAATTTGAGAGCAGTCTTCCAAATTTATATCAAGGACCTGGAGAAGTACTCA 3013
QY 4865 TGATAGCAGGAGAGAGCGGCACTGTGCTTTGAGAGCTGAAGAAAGTGAACAGTCCC 4924
Db 3014 TGATAGCAGGAGAGAGCGGCACTGTGCTTTGAGAGCTGAAGAAAGTGAACAGTCCC 3073
QY 4925 TGGCCAGTCCCACTGCTGCCAGCCGACATCTCAACCAACATTTTGAAGCTGTCA 4984
Db 3074 TGGCCAGTCCCACTGCTGCCAGCCGACATCTCAACCAACATTTTGAAGCTGTCA 3133
QY 4985 AGGGCTGCCACTTTGTTGGGCGAGGCAAGATGAGAAACGGGCTCTGATCTGTGCAAGCA 5044
Db 3134 AGGGCTGCCACTTTGTTGGGCGAGGCAAGATGAGAAACGGGCTCTGATCTGTGCAAGCA 3193
QY 5045 TGCCAGCAAAAGTCTCATTTCTCGCTACACGAAACCTCAGCAAAATACTGCAATCCGGA 5104
Db 3194 TGCCAGCAAAAGTCTCATTTCTCGCTACACGAAACCTCAGCAAAATACTGCAATCCGGA 3253
QY 5105 AAGAGATAGACACTCAGAGCCCTSCAGCTGTATCCACTTCAACAAATTACAGTATCCTCA 5164
Db 3254 AAGAGATAGACACTCAGAGCCCTSCAGCTGTATCCACTTCAACAAATTACAGTATCCTCA 3313
QY 5165 TTGAAACCAATAATTTCTACGAATTCGACATGAAGAGTACACGCTCGAGGAATTCCTGG 5224
Db 3314 TTGAAACCAATAATTTCTACGAATTCGACATGAAGAGTACACGCTCGAGGAATTCCTGG 3373
QY 5225 ATAAGATGACCAATCTTGGCAGCTGCTGTGTTGGCTCTTCCAAACAGCTTCCCTG 5284
Db 3374 ATAAGATGACCAATCTTGGCAGCTGCTGTGTTGGCTCTTCCAAACAGCTTCCCTG 3433
QY 5285 TCTCAATCTGTGAGTGAAACAGCGCAGGCGAGGAGGAGTACTTGTCTGTGTTCCAGC 5344
Db 3434 TCTCAATCTGTGAGTGAAACAGCGCAGGCGAGGAGGAGTACTTGTCTGTGTTCCAGC 3493
QY 5345 AATTGGAGTGTTCGTGGATTCTTACGGAAGACGTAGCCGACAGACGATCTCAAGTGGGA 5404
Db 3494 AATTGGAGTGTTCGTGGATTCTTACGGAAGACGTAGCCGACAGACGATCTCAAGTGGGA 3553
QY 5405 GTCCTTACCTTTGGCCTTTGCCCTTACAGAGAACCTTATCTGTTGTGACCACTTCAACT 5464
Db 3554 GTCCTTACCTTTGGCCTTTGCCCTTACAGAGAACCTTATCTGTTGTGACCACTTCAACT 3613
QY 5465 CACTCGAAGTAATTTAGATCCAGCAGCTCCTCAGCAGGAGCCCTGCGCCGAGCGTACC 5524
Db 3614 CACTCGAAGTAATTTAGATCCAGCAGCTCCTCAGCAGGAGAACCTTGTGCGCCGAGCGTACC 3673
QY 5525 TGGACATCCCGAAACCCCGCTTACCTGGGCGCTGCCATTTCTCAGGAGCGATTACTTGG 5584
Db 3674 TGGACATCCCGAAACCCCGCTTACCTGGGCGCTGCCATTTCTCAGGAGCGATTACTTGG 3733
QY 5585 CGTCTCATACCAGGATAAATTAAGGTCAATTTGCTGCAAGGAAACCTCGTGAAGAGT 5644
Db 3734 CGTCTCATACCAGGATAAATTAAGGTCAATTTGCTGCAAGGAAACCTCGTGAAGAGT 3793
QY 5645 CCGGCACTGAACACCAACCGGGGCCCTGCCACTCCGACAGCAGCCCAACAGCGAGGCC 5704
Db 3794 CCGGCACTGAACACCAACCGGGGCCCTGCCACTCCGACAGCAGCCCAACAGCGAGGCC 3853
QY 5705 CACCCAGTGAACAGGACATCAACAGCGCTGGCTCCAGCCAGCGCCGCCCGGAAG 5764

1636 AGCCGAGGCTCAAGAAATCAAGAGCAGGAGTACCAGGCTCAAGTGAAGAAATGAGG 1695
1936 AGCCGAGGCTCAAGAAATCAAGAGCAGGAGTACCAGGCTCAAGTGAAGAAATGAGG 1877
1696 TTGATGATGAATCAGTTGGAAGAGTCTTTGTCTCAGCAAGAGAGCGAGTGTCTCTAC 1755
1876 TTGATGATGAATCAGTTGGAAGAGTCTTTGTCTCAGCAAGAGAGCGAGTGTCTCTAC 1817
1756 GAATCTGAGCTGAGAGAGTCTGGCTTGTCTGCTGAAGAAATCAAGCGGAAGCGACAGNA 1815
1816 GAATCTGAGCTGAGAGAGTCTGGCTTGTCTGCTGAAGAAATCAAGCGGAAGCGACAGNA 1757
1816 TGTTCAGATAAATCTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAAGTGGGAGAAATGCG 1875
1756 TGTTCAGATAAATCTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAAGTGGGAGAAATGCG 1697
1876 AAATCTGGAAGATCAATGCTGAGCAGAGCTCAAAATTCAGAGAGCTCCAGAGAAATCTG 1935
1696 AAATCTGGAAGATCAATGCTGAGCAGAGCTCAAAATTCAGAGAGCTCCAGAGAAATCTG 1637
1936 GAGAAGGCT-----GCA 1947
1636 GAGAAGGCTGTAAGAGCCAGCAGGAGGCCACCGAGCTGCTGCAGNAATATCCGCCAGCA 1577
1948 AAGGAGCGAGCCGAGAGGAGTGGAGAAAGCTGCAGAACCGAGAGGATCTTTCTGAAGGC 2007
1576 AAGGAGCGAGCCGAGAGGAGTGGAGAAAGCTGCAGAACCGAGAGGATCTTTCTGAAGGC 1517
2008 ATCAGAAAGAGCTGTTGAGCTGAGAGCTGAGGAGCCGCAATCTCTGAGAGCAAGGTAAAG 2067
1516 ATCAGAAAGAGCTGTTGAGCTGAGAGCTGAGGAGCCGCAATCTCTGAGAGCAAGGTAAAG 1457
2068 AGACTAGAGACCATGGAGCTTAGAGAAAACAGACTGAAGGATGACATCCAGACAAAATCC 2127
1456 AGACTAGAGACCATGGAGCTTAGAGAAAACAGACTGAAGGATGACATCCAGACAAAATCC 1397
2128 CAACAGATCAGCAGATGGCTGATAAAATTTCTGGAGCTGAGAGAAACATCGGAGGCC 2187
1396 CAACAGATCAGCAGATGGCTGATAAAATTTCTGGAGCTGAGAGAAACATCGGAGGCC 1337
2188 CAAGTCTCAGCCAGCACCTAGAAGTGCACCTGAAAACAGAGAGCAGCACTATGA-GGA 2246
1336 CAAGTCTCAGCCAGCACCTAGAAGTGCACCTGAAAACAGAGAGCAGCACTATGAGGA 1277
2247 AAAGATTAAGTGTGGACAAATCAGATAAAGAAAGACCTGGGTGACAGAGAGACACTGGA 2306
1276 AAAGATTAAGTGTGGACAAATCAGATAAAGAAAGACCTGGGTGACAGAGAGACACTGGA 1217
2307 GAACATGATGCAGACACAGAGGAGGAGGCCATGAGAGGCGCAAAATTTCTCAGCGAACA 2366
1216 GAACATGATGCAGACACAGAGGAGGAGGCCATGAGAGGCGCAAAATTTCTCAGCGAACA 1157
2367 GAAGGGGATGATCAATGCTATGATTCAGAGTCAAGATTCAGATCCCTGGAACAGAGGATTTGGA 2426
1156 GAAGGGGATGATCAATGCTATGATTCAGAGTCAAGATTCAGATCCCTGGAACAGAGGATTTGGA 1097
2427 ACTGTCGAGCCAAATAAATTCAGCAATAGCAGTCTTTTACCCAAAGNACATGAA 2486
1096 ACTGTCGAGCCAAATAAATTCAGCAATAGCAGTCTTTTACCCAAAGNACATGAA 1037
2487 GGCCCAAGAGAGATGATTTCTGAATCTCAGGCCAACAGAAAATTTTACCTGGAGACACAGGC 2546
1036 GGCCCAAGAGAGATGATTTCTGAATCTCAGGCCAACAGAAAATTTTACCTGGAGACACAGGC 977
2547 TGGGAAGTTGGAGGCCAGAACCGAAAACCTGGAGGAGAGCTGGAGAGATCAGCCACCA 2606
976 TGGGAAGTTGGAGGCCAGAACCGAAAACCTGGAGGAGAGCTGGAGAGATCAGCCACCA 917
2607 AGACCAGAGTGAAGAAATCGGCTGCTGGAATCTGGAGACAAAGATTCGGGAGGTCACTCT 2666
916 AGACCAGAGTGAAGAAATCGGCTGCTGGAATCTGGAGACAAAGATTCGGGAGGTCACTCT 857
2667 AGACACGAGGAGCAAAAATCGGAGTCAAGCGCCAGCTCAGACAGCTACAGCTTCCCT 2726

856 AGACACAGGAGCAGAAAATCGAGCTCAAGCGCCAGCTACAGAGCTACAGCTCTCCCT 797
2727 GCAGAGCGGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGCGGCCCTTGAGAGCCA 2786
796 GCAGAGCGGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGCGGCCCTTGAGAGCCA 737
2787 GCTTCGCGCAGCGAAGACAGAGCTGGAAGAGACACACAGCAGAAAGCTGAAGAGAGATCCA 2846
736 GCTTCGCGCAGCGAAGACAGAGCTGGAAGAGACACACAGCAGAAAGCGAAGAGAGATCCA 677
2847 GGCACTCAOCCGACATAGAGATGAATTCAGCGCAAAATTTGATGCTCTTCTGAAGCTG 2906
676 GGCACTCAOCCGACATAGAGATGAATTCAGCGCAAAATTTGATGCTCTTCTGAAGCTG 617
2907 TACTGTATCAACAGACCTGGAGGAGCAGTAAACAGCTCACCAGGACCAACGCTGAAT 2966
616 TACTGTATCAACAGACCTGGAGGAGCAGTAAACAGCTCACCAGGACCAACGCTGAAT 557
2967 CAACAACCAAAATCTTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGACGA 3026
556 CAACAACCAAAATCTTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGACGA 497
3027 GATTGTAACTCGAAGTGAAGTGAACCATCTCCGCCGGGAGATCAACGGAACGAGAGAT 3086
496 GATTGTAACTCGAAGTGAAGTGAACCATCTCCGCCGGGAGATCAACGGAACGAGAGAT 437
3087 GCAGCTTACCAGCCAGAGCAAAACGATGAGGCTTCTGAAGACCAACGCTGCTGGA 3146
436 GCAGCTTACCAGCCAGAGCAAAACGATGAGGCTTCTGAAGACCAACGCTGCTGGA 377
3147 GGAACAGGTCATGATTTGAGGCTTAAACGATGAGCTGCTAGAAAAAGAGCGGAGTG 3206
376 GGAACAGGTCATGATTTGAGGCTTAAACGATGAGCTGCTAGAAAAAGAGCGGAGTG 317
3207 GGAGGCTTGGAGGAGCTCTCTGGGTGATGAGAAATCCAGTTTGAGTGTCCGGTTCGAGA 3266
316 GGAGGCTTGGAGGAGCTCTCTGGGTGATGAGAAATCCAGTTTGAGTGTCCGGTTCGAGA 257
3267 GCTCCAGAGATGCTGACACCCGAGAAAACAGAGCAGGCGGAGAGCCGATCAGCGGATCAC 3326
256 GCTCCAGAGATGCTGACACCCGAGAAAACAGAGCAGGCGGAGAGCCGATCAGCGGATCAC 197
3327 CGAGTCTGCCAGTGTGGAGCTGGCAGTGAAGGACCAAGCTGAGATTCTTCGCTCT 3386
196 CGAGTCTGCCAGTGTGGAGCTGGCAGTGAAGGACCAAGCTGAGATTCTTCGCTCT 137
3387 GCAGCAGGCTCTCAAGAGCAGAAAGCTGAAGCGCGGAGAGCTCTCTGACAAAGCTCAATGA 3446
136 GCAGCAGGCTCTCAAGAGCAGAAAGCTGAAGCGCGGAGAGCTCTCTGACAAAGCTCAATGA 77
3447 CCTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGA 3506
76 CCTGGAGA-----AGAGCTGGA 59
3507 GACTGAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGCAG 3564
58 GACTGAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGCAG 1

RESULT 15

ABA08361

ID ABA08361 standard; cDNA; 3131 BP.

XX ABA08361;

XX 11-JAN-2002 (first entry)

XX Human RHO/RAC effector homologue-encoding cDNA, SEQ ID NO:137.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;

XX haematopoiesis regulation; tissue growth; immunomodulator; activin;

XX inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;

KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; Proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antilucer; ss.

XX Homo sapiens.

OS WO200157188-A2.

PN 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US003800.

XX 03-FEB-2000; 2000US-00496914.

PR 27-FEB-2000; 2000US-00506875.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

DR P-PSDB; ABB11117.

PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.

XX Claim 1; Page 387-388; 1963pp; English.

PS Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention

XX

SQ Sequence 3131 BP; 861 A; 835 C; 806 G; 629 T; 0 U; 0 Other;
Query March 41.6%; Score 2564.8; DB 4; Length 3131;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2781; Conservative 0; Mismatches 2; Indels 195; Gaps 2;
QY 2857 GCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATC 2916
DB 1 GCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATC 60
QY 2917 ACAGACCTGGAGGAGCAGCTAAACCCAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAA 2976
DB 61 ACAGACCTGGAGGAGCAGCTAAACCCAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAA 120
QY 2977 AACCTTCTACTTGTCCAAAACAACTCGATGAGGCTTCTGCGCCCAACGACGAGATTGTACAA 3036
DB 121 AACCTTCTACTTGTCCAAAACAACTCGATGAGGCTTCTGCGCCCAACGACGAGATTGTACAA 180
QY 3037 CTGCGAAGTGAAGTGGACCATCTCCCGCCGGGAGATCAACGAAACGAGAGATGACGCTTACC 3096
DB 181 CTGCGAAGTGAAGTGGACCATCTCCCGCCGGGAGATCAACGAAACGAGAGATGACGCTTACC 240
QY 3097 AGCCAGAGCAACGATGGAGGCTCTGAAGACCAACGTCACCATGCTGGAGGAAACAGGTC 3156
DB 241 AGCCAGAGCAACGATGGAGGCTCTGAAGACCAACGTCACCATGCTGGAGGAAACAGGTC 300
QY 3157 ATGGATTGGAGGCCCTAAACGATGAGTCTCTAGAAAAAGAGCGGAGTGGAGGCGCTGG 3216
DB 301 ATGGATTGGAGGCCCTAAACGATGAGTCTCTAGAAAAAGAGCGGAGTGGAGGCGCTGG 360
QY 3217 AGGAGCGTCTGGGTGATGAGAAATCCAGTTTGGAGTGTGGGTTCGAGAGCTGCAGAGA 3276
DB 361 AGGAGCGTCTGGGTGATGAGAAATCCAGTTTGGAGTGTGGGTTCGAGAGCTGCAGAGA 420
QY 3277 ATGCTGACACCCGAGAAAACAGCAGCGGCGAGAGCCGATCAGCGGATCACCAGTCTCGC 3336
DB 421 ATGCTGACACCCGAGAAAACAGCAGCGGCGAGAGCCGATCAGCGGATCACCAGTCTCGC 480
QY 3337 CAGTGTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCTCGCTTCGACAGAGCT 3396
DB 481 CAGTGTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCTCGCTTCGACAGAGCT 540
QY 3397 CTCAGAGCAGAGAGCTGAAGGCGGAGAGCCTCTCTCACAAGCTCAATGACCTGGAGAAG 3456
DB 541 CTCAGAGCAGAGAGCTGAAGGCGGAGAGCCTCTCTCACAAGCTCAATGACCTGGAGAAG 600
QY 3457 AAGCATGCTATGCTTGAATGAATGCCGGAAGCTTACAGCAGAAAGCTGGAGACTGAACGA 3516
DB 601 AAGCATGCTATGCTTGAATGAATGCCGGAAGCTTACAGCAGAAAGCTGGAGACTGAACGA 660
QY 3517 GAGCTCAACAGAGGCTCTGGAAGCAAGCCAAATACAGCAGCAGATGGACCTGCAG 3576
DB 661 GAGCTCAACAGAGGCTCTGGAAGCAAGCCAAATACAGCAGCAGATGGACCTGCAG 720
QY 3577 AAAAATCACATTTTCCGCTGACCTCAAGGACTGCAAGAAGCTCTAGATCGGGCTGATCTA 3636
DB 721 AAAAATCACATTTTCCGCTGACCTCAAGGACTGCAAGAAGCTCTAGATCGGGCTGATCTA 780
QY 3637 CTGAAGCAGAAAGAGTACTTGGAGTATCAGCTGGAAGAAACATTGAGTTCTTATCT 3696
DB 781 CTGAAGCAGAAAGAGTACTTGGAGTATCAGCTGGAAGAAACATTGAGTTCTTATCT 840
QY 3697 CATGAAAGGTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCATTGATTTCTG 3756
DB 841 CATGAAAGGTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCATTGATTTCTG 900
QY 3757 CAAGCCAAAATGGAACCACTGCTTAAAGAAAGAGGTTTATTAGTCGACGGAAGAG 3816
DB 901 CAAGCCAAAATGGAACCACTGCTTAAAGAAAGAGGTTTATTAGTCGACGGAAGAG 934
QY 3817 GACCTGCTTTACCCACACAGGTTCTCTGCAATGAGTGAAGCTGGCCCTGGAG 3876
DB 935 -----AGGTTCTCTGCAATGAGTGAAGCTGGCCCTGGAG 975

